IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-5, Issue-11 E-ISSN: 2347-2693

Web Recommendation Using Microblogging Information

U. Lakshmi Prasanna[1*], A. Revathi[2]

[1*]Dept. of IT, VNR Vignana Jyothi Institute of Engineering and Technology, JNTUH, Hyderabad, India
[2]Dept. of IT, VNR Vignana Jyothi Institute of Engineering and Technology, JNTUH, Hyderabad, India

Corresponding Author: prasu.lucky01@gmail.com

Available online at: www.ijcseonline.org

Received: 02/Oct/2017, Revised: 16/Oct/2017, Accepted: 10/Nov/2017, Published: 30/Nov/2017

Abstract- As of late, the gap between online business and person to person communication has turned out to be m ore and more indistinct. Numerous online businesses reinforce the social sign in system where customers can sign-in on their portals by applying their informal organization characters, like their Twitter or Facebook IDs. Users additionally can announce their recently bought items on social networking or microblogs by mentioning the corresponding product url from the online business sites. In this paper, we put forward an innovative solution for “cross site cold start web product recommendation” to endorse different items from “e-commerce” sites for users at “microblogging or social networking” sites in “cold start positions”, a very rare concept explored before. The foremost task is to utilize the information fetched from microblogging or social interacting. In exact, we propose learning the two client and items element depictions called client embeddings and item embeddings respectively from the information fetched through online commercial sites using “recurrent neural systems”. And later applying “Altered Gradient boosting trees” model to convert user long range informal statement keywords into user embedding’s. We at that point build up a component based “Lattice factorization method” which can be used to learn client embedding’s for “cold-start product recommendation”.

Keywords- E-commerce, recurrent neural networks, demographic, microblogs, product recommendation.

I. INTRODUCTION

As Today’s creation is becoming fully connected through the Internet and Internet is the source for the most needed information. The internet generates the huge size of information every day. “E-commerce” sites like the Amazon, eBay, Flipkart and Alibaba features many of the social network characteristics such as views, updates and the interface between product suppliers and purchasers. Now a day’s many “e-commerce” sites integrated the ease of microblogging sign up, to allow new customers to login with their prevailing accounts from microblogging or “social networking” sites namely Google Plus, Twitter and Facebook. Twitter and Facebook presented a new option last year, which had attracted and allowed number of customers to purchase products quickly from their web portal by using a “buy” key option for ordering items based on some advertisements. In this Recommendation shows a key part in many areas and has appealed a lot of investigative interest. Here we think about an exciting concept of endorsing products from “e-commerce” sites to customers at “social networking” destinations without any verifiable purchase histories, i.e., in "cold start" circumstances. We termed this issue as “cross-site cold-start product recommendation”. Even though online based product endorsing has been widely considered previously [1], [4], [5], many of those examinations highly focused on developing product endorsements algorithms at “e-commerce” sites using clients' verifiable purchase history. To the top of our insight, “cross-site cold-start product recommendation” has been seldom contemplated sometimes of lately.

Flipkart, Amazon, Netflix and e-bay can browse and choose items that they are interested in, the poster also plays a key role in the system to recommend the product to users. Then the objects that the system thought as a preeminent one will be the preferred match to the product for the recommendation. Afterward, the customer may provide feedback (such as rating, usually represented as a score between, for example, 1 and 5, also the reviews make a huge decision in the product buying) on how the users think about an item after she/he has felt the item. One main task of the recommendation system is to understand the users’ personalized preferences from their historic rating behaviors.

In this study, we focused on the significant problem of endorsing objects from “e-commerce” site to “social networking” site users without any past purchase histories, i.e., in “cold-start” conditions [10]. Improving the recommendation precision is the challenging task for new (or rarely rated) products and for the new (or inactive) users. Relating to the common items, for the latest released ones and the old items that are not often rated by users, it is hard for standard recommendation approaches such as the common filtering approach to offer high-quality approvals.

II. MICROBLOGGING ATTRIBUTES

Here we focused on extracting substantial information from the microblogging rich customer.

Demographic Attributes

Demographic profile is very significant in advertising and mostly in product acceptance. User's material such as gender, age and education, career and interest. Demographic attributes became very significant in product promotion, particularly in product acceptance for buyers.

Text Attributes

Text Attributes reflect user's ideas and interests about certain subjects. Topic sharing for each customer will be obtained after combining all the information from microblogs. In word embedding's, every dimension signifies underlying features of the word and interpretation of related words are near in the dormant space. "Skip-Gram model" is used to understand words distributed representation and "word vectors" of all the indications in a user's distributed document aggregated as the "user's embedding vector".

Network Attributes

In Internet based "social networking", it is witnessed that, users allied with each other (e.g., Liking same pages) are probable segmenting related ideas and interests. we propose subject models for learning hidden groups of interests and followings. All the followings and similar interests of a user is aggregated as an individual document. Each user is treated as token. By this way, we should be able to extract latent groups of users who shares similar interests and ideas (called "following topics"), and each user is represented as a first choice distribution for hidden groups.

Temporal Attributes

Up to some extent lifestyles of the "social media or microblogging" users are reflected by temporal activity outlines. In exact, there might be a connection among temporal activities outlines and users' buying preferences. We considered two categories of temporal action distributions, called as daily movement distributions and weekly action distributions for item recommendation.

III. SYSTEM ARCHITECTURE

In the proposed method we used the connected users across microblogging places and "e-commerce" sites (clients who have microblogging IDs and have purchased products on "e-commerce" sites) shown in Fig.1.

Social Media

On social media website we have social interacting information such by means of group media likes and we study about how to extract information from microblogging from the rich user. From this microblogging feature depiction can be created.

In classifier training model microblogging feature choice created on demographic attributes, text attributes, network attributes and temporal attributes.

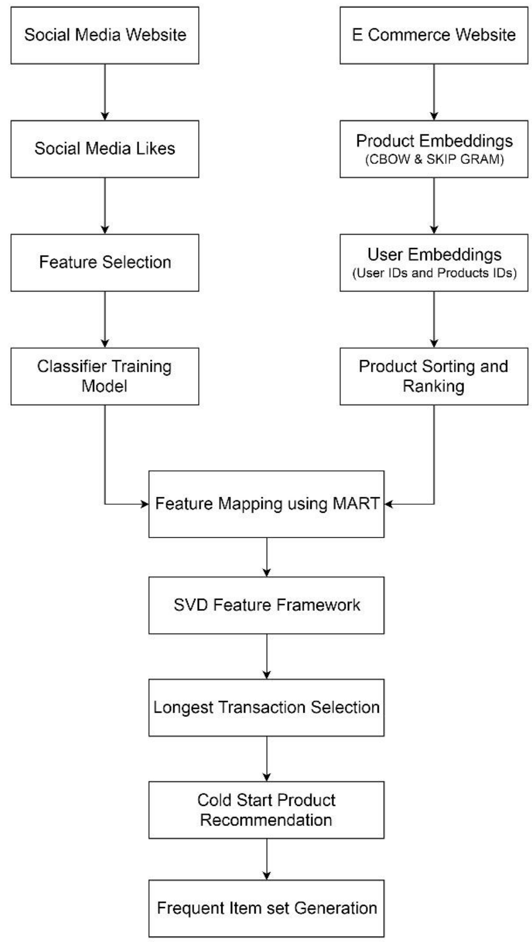

Fig.1. System Architecture

E Commerce

We considered two "recurrent neural" models to train product embedding's, the "CBOW Model" and the "Skip-Gram Model". "CBOW Model" guesses the present product using the proximate context and Skip Gram archetypal guesses similar products.

The buying history of a customer measured as sentence which has an order of different product ID's "word tokens". Customer unique ID is appended at the start of every sentence. In Learning process customer IDs and products IDs are considered as word indications. By default customer ID is associated with respective customer purchase records. Element descriptions will be generated by making use of the data from every customer on the "e-commerce" site through intense learning.

SVD Feature Framework

SVD Feature is framed on the basis of the customary framework factorization approach. It considers factorization in three perspectives, to be specific Global components (likewise called as dyadic elements), consumer components and object highlights.

The records collected from "E-commerce" site and social media features are mapped using MART. By using singular value decomposition we can select the longest transaction. A common item set will be generated by SVD from the longest transaction and finally we represents items which are generating frequently.

IV. PRODUCT RECOMMENDATION USING SVD FEATURE FRAMEWORK:

Global components, customer components and object highlights are three viewpoints considered for the SVD Feature framework factorization.

$$\hat{r}_{u,p}(A^{(u)}, B^{(p)}, C^{(u,p)}) = \mu + \sum_j x_j^{(G)} C_j^{(u,p)} + \sum_j x_j^{(U)} A_j^{(u)} + \sum_j x_j^{(P)} B_j^{(p)} + \left(\sum_j A_j^{(u)} M_j\right)^T \left(\sum_j B_j^{(p)} N_j\right)$$

Where $A^{(u)}, B^{(p)}, C^{(u,p)}$ are Global components, customer components and object highlights correspondingly. The two "latent vectors" M_j and N_j holds the jth customer feature and item feature respectively. Here $x_j^{(G)}$, $x_j^{(U)}$, $x_j^{(P)}$ are worldwide, customer and item bias parameters respectively. The records collected from "E-commerce" site and social media features are mapped using MART.

MART will be used as a channel to map the user's social interacting structures to other feature representation for item endorsement. In precise, we advise knowledge of both user's and product's feature signs (called as customer embedding's and item embedding's) from the data obtained through "e-commerce" sites by using "recurrent neural networks" and later by applying "gradient boosting trees method" for converting user's microblogging features into customer embedding's. Further we develop a factorization method which can support and improvise the customer embedding's for "cold start cross site" product recommendation.

V. RELATED WORK

Recommender frameworks: Lately, "The Grid factorization approach" [6] attracted a great number of research interests. Many research examinations concentrated on consolidating helper data [1], [7], [8], into the "Lattice Factorization" approach due to the huge increase in size of web data. Two such research investigations are the "SVD Feature" and the "Factorization Machine".

We considered the initial data mining research on "social networking" sites. My effort is likewise identified with thinks about on "programmed client profiling" and "cross-site" association deduction. My effort is on the basis of these examinations, predominantly on the lines of "cross-site cold-start" suggestions. Despite the detail that having a few uniformities, we are managing a certain assignment of exceedingly functional regard, "cold-start" product suggestion to microblogging end users. Main investigations are from [9] by relating clients crosswise over eBay and Facebook. On the other way, they focused mainly on the product classification level and the product brand level buying interests based on a "trained classifier", which can't be straightforwardly associated to our "cross-site cold-start" item suggestion assignment.

In addition earlier studies features just considered age, gender and few social media attributes such as the likes in Facebook and didn't included a comprehensive variety of highlights as considered in our approach. In conclusion, they didn't study and consider the way to handle multi-dimensional variety data from the internet created by "social media" sites into a ready to embed structured data on the "e-commerce" site, a way to address the "cross-site cold-start" situation.

Cross-Site suggestion: One important strategy for "cross-site" suggestion is to Transfer Learnings. The concept is to study and absorb the "transferable knowledge" from one system to apply the same in another system known as source and target system respectively. G. J. Gordon and A. P. Singh [2] proposed "Collective Matrix factorization" method for estimating the associations of different objects. While allocating parameters in dormant, this can be achieved by factorizing multiple matrices concurrently. Li [3] made an effort to exchange client item rating outlines from a secondary or supporting matrix in another area to the objective space through Code books.

VI. IMPLEMENTATION

The data flow model consist of various steps as shown in Fig.2.

User Login: In "social networking" sites user login is a form where user needs to enter his login credentials password and user name. User profile will be opened successfully when provided login IDs are correct, otherwise shows invalid login details. It is designed for easy logins to the end users and also to provide more consistent demographic data to web developers when user registers with his details.

Once user profile is opened we can see user's menu with search friends, view friend requests, search posts, recommended to friend and purchase, user search history

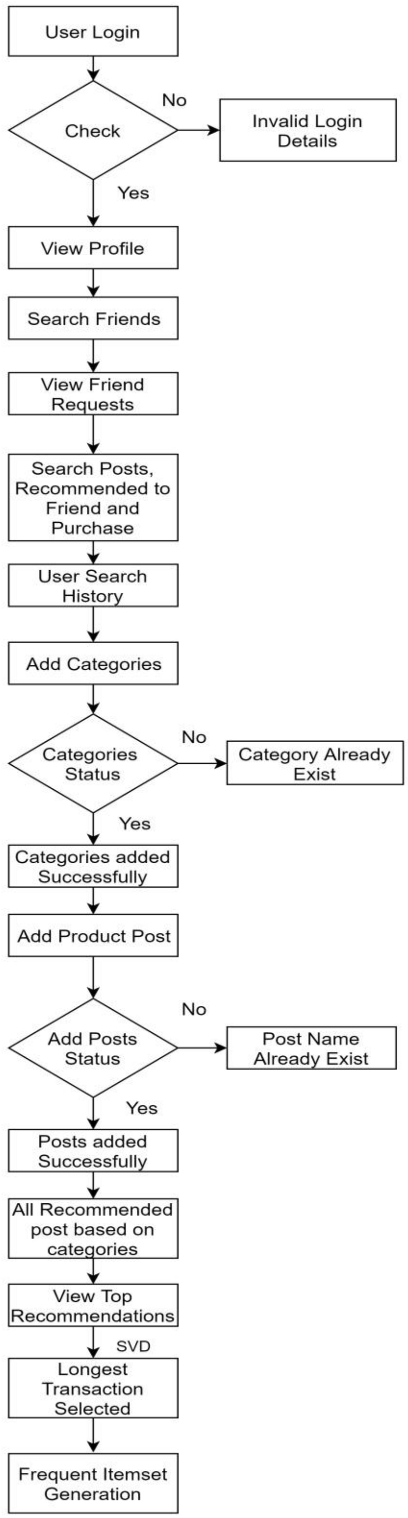

Fig.2. Data flow model

View Profile: The administrator can view the registered social media users in this module. The profile of the user along with the purchase history can be viewed here.

Search Friends: In this module the users can find the friends they are looking for in the social media if they have registered in the same media.

View Friend Request: Other user's friends requests can be viewed in this module. In this module the user can either accept or deny the request given by them.

User Search History: The previous search made by the registered users can be viewed by the administrator in this module.

Add categories: The structure and names of the "e-commerce" categories that you can store Similar to the physical store, the basic objective for the online store is to make it easy for customers to search and purchase items. One can define the unlimited number of "e-commerce" product categories and subcategories. New categories can be added if it not previously exist.

Add product post: Each product post should include the category, post name, price, description, images. Post will be added successfully only when if it is not already exist

Product recommendation: Basically when the records collected from "e-commerce" website and social media features are mapped using MART. An array of current post categories will be generated. Categories array will used to query the database to get other items within these categories. We training a significant snag of endorsing items from the "e-commerce" sites to the users of "social networking" with no product buying histories, namely in "cold-start" positions by using the identified six key demographic qualities i.e, the age, Gender, Marital status, Interest, Education, and Career.

Longest transaction: The records collected from "E-commerce" site and social media features are mapped using MART. By using singular value decomposition we can select longest transaction

Frequent Item Set: The items which are frequently purchased by costumers are finally generated list in this module.

VII. EXPERIMENTS

Our task needs information from "e-commerce" site and "social networking" site. We utilized an "e-commerce" dataset with the large volume of data. "Tree-based methods" offers added feasibility to learn and understand the qualified status of each quality. We derive the statistic of the qualified meaning of the respective feature for MART established on top of the training data. In MART, a quality value will be represented by each section. Initially, all the "regression trees" will be examined. Compute for respective element support towards the "Cost function" by summing-up all the values contributed by all the divided nodes by this feature. We outline the feature impact towards the damage function in reducing the squared error. We can add up each values contributed by its possible attributes as its overall contribution for the respective attribute.

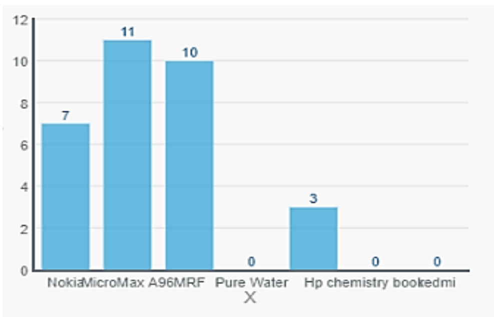

Fig.3. Relative product importance ranking

We will classify users into two different sets called as the test set and the training set for evaluation. We will sample negative products and positive products on the proportion of 1:1 for every single user in the training set which means we will be having the equal number of positive and negative products. We will sample negative and positive products randomly on the proportion of 1:50 for every single user from the same category of the product in the test set which means we will be having 50 negative products for each positive product. We first assess the effectiveness of product endorsement on D_{dense}, where d is the percent of related users, used for the training data, and the left over percent linked users are used as experimental test records. To study the effectiveness by varying total of training data with relative products importance ranking shown in Fig.3. the highest ranking is the positive product and lowest ranking sown is the negative product.

Table.1. Frequent itemset list

S.No	Items
1	Footballs
2	Diesel fuel
3	CD Player
4	Oil Filters
5	Floor Wax
6	Paint Rollers
7	Shampoo
8	Trash Bags
9	Caulking
10	Fishing Rods
11	Shoe Polish
12	Motorcycle Helmet
13	Antiseptics
14	Rubbing Alcohol
15	Sports Car Bodies
16	Fertilizers
17	Skis
18	Linoleum
19	Ballpoint Pens
20	Tool Boxes
21	Polish Remover

The records collected from "E-commerce" website and social media features are mapped using MART. By using singular value decomposition we can select the longest transaction. Common items set will be generated by SVD from the longest transaction. In the above table we represents items which are generating frequently.

VIII. CONCLUSION

In this paper, we attempted to solve an innovative issue, "cross-site cold-start product recommendation", i.e., prescribing items to microblogging users from "e-commerce" sites without having previous purchase transactions. In the "e-commerce" sites, Products and the microblogging customer IDs will be denoted on the common underlying feature space over the feature learning with "recurrent neural networks". "Feature mapping functions" developed using "modified gradient boosting trees method" which relates user's features obtained from microblogging sites into feature representations learned from "e-commerce" sites. This can be done by utilizing a set of related users over "e-commerce" sites and microblogging sites as a connection.

The related client highlights can be successfully joined into an element based framework factorization methodology for "cold-start product recommendation". The outcome demonstrates that our planned structure is for sure successful in addressing the "cross-site cold-start product recommendation" issue. We reliance on this study that it will be having great impact on both research and industry groups. Later on, further advanced modern deep learning methods, like "Convolutional Neural Networks" can be researched for learning of different user features. We will similarly study the feasibility of enriching the current "feature mapping technique" considering industry experts ideas in the process of continuous learning.

REFERENCES

[1] B. Xiao as well as I. Benbasat, "*Web based business item proposal operators: Use, attributes, and effect*" MIS Quarterly, 2007,vol. 31, pp. 137 to 209.

[2] J. Gordon and A. P. Singh, "*Social knowledge through aggregate network factorization,"* in Proceeding's fourteenth ACM SIGKDD international conference knowledge disclosure data mining, pp. 650 to 658, 2008.

[3] B.Li, X. Xue, and Q. Yang "*Can pictures and records collaborate? Cross-domain collaborative clarifying for sparsity lessening*," in Proceedings twenty first international joint Conference artificial intelligence pp. 2052 to 2057, 2009.

[4] M. Gearing, "*Trade deals expectation and thing suggestions utilizing client socioeconomics at stock level*," SIGKDD Exploration Newsletter, Dec. 2008. vol. 10, no. 2, pp. 84-89.

[5] Y. Zhang and J. Wang, "*Chance demonstrate for E-trade suggestion: Correct item; perfect period*," in Processes thirty sixth Institute, ACM SIGIR Conference research, Create Information Recovery, pp. 303 to 312, 2013

[6] Y. Korean, C. Volinsky, "*Network factorization strategies for recommender frameworks,*" Computer, volume 42, number 8, pp. 30 to 37, August 2009.

[7] L. Hong, B. Davison, and A. S. Doumith, "*Co-factorization machines: Modelling client interest and foreseeing singular choices in Twitter,"* in Proc. 6th ACM International. Conf. Web Hunt Data Mining, pp. 557 to 566, 2013.
[8] M. R. Lyu, H. Ma, I. king, and T. C. Zhou, "*Enhancing recommender frameworks through consolidating public logical data*," ACM Transactional. Information. System, no. 2, 2011.vol. 29.
[9] M. Pennacchiotti and Y. Zhang, "*Recommending branded items from social interacting sites,"* in Proceeding's seventh ACM Conference. Recommender Systems. Hong Kong, China, pp. 77 to 84, October 12 to 16, 2013.
[10] Wayne Xin Zhao, Edward y.chang , Ji-oRong Wen, "*connecting social media to e-commerce cold start product recommendation using microblogging information*" volume 28, number 5, may 2016.