ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Connecting Social Media to E-Commerce Site Using Cold Start Product Recommendation

Vedwati V Jadhav[1] , Pooja P Ramane[2], Shraddha A Shelar[3],Snehal S Vahile[4] ,Poonam B Thorat[5], Prema B Sahane[6]

[1,2,3,4]Department of Computer Engineering, P.K. Technical Campus, Chakan, Pune, India

[5]Assistant Professor, Department of Computer Engineering, P.K. Technical Campus, Chakan, Pune, India

[6]Professor &HOD, Department of Computer Engineering, P.K. Technical Campus, Chakan, Pune, India

ABSTRACT: In recent years, the boundaries between e-commerce and social networking became progressively blurred. Several e-commerce websites support the mechanism of social login wherever users will sign in the websites victimization their social network identities like their Facebook or Twitter accounts. Users also can post their fresh purchased product on microblogs with links to the e-commerce product websites. During this paper we have a tendency to propose a unique answer for cross-site cold-start product recommendation that aims to advocate product from e-commerce websites to users at social networking sites in "coldstart" things, a haul that has seldom been explored before. A serious challenge is the way to leverage data extracted from social networking sites for cross-site cold-start product recommendation. We propose to use the coupled users across social networking sites and e-commerce websites (users United Nations agency have social networking accounts and have created purchases on e-commerce websites) as a bridge to map users' social networking options to a different feature illustration for product recommendation. In specific, we have a tendency to propose learning each users' and merchandises' feature representations (called user embeddings and product embeddings, respectively) from information collected from e-commerce websites victimization continual neural networks so apply a changed gradient boosting trees methodology to remodel users' social networking options into user embeddings. We have a tendency to then develop a feature-based matrix factorisation approach which might leverage the learnt user embeddings for cold-start product recommendation. Experimental results on an oversized dataset made from the biggest Chinese microblogging service SINA WEIBO and also the largest Chinese B2C e-commerce web site JINGDONG have shown the effectiveness of our planned framework.

KEYWORDS: E-commerce, Product Recommender, Product Demographic, Microblogs, Recurrent Neural Networks.

I. INTRODUCTION

Nowadays, Recommender Systems, aiming at serving to users realize relevant and attention-grabbing things from the knowledge era, are wide studied and applied in varied fields starting from e-commerce to medication prediction .Besides the enumerable studies on rising the advice performance the way to fittingly justify their commendation results and ultimately persuade users to simply accept them is additionally an awesome challenge in each analysis and engineering fields. Though several novel algorithms have well-tried that they need achieved smart, even wonderful performance in varied matrices on offline datasets, feedbacks from on-line applications show that users wouldn't invariably trust and follow the machine-produced results, that in additional hinders its wider development in real society Recently, the acquisition intention of users has attracted abundant attention from scientific community. Completely different from ancient recommender systems, they specialize in finding the factors which might verify one's temperament to buy merchandise on-line. In fact, the $64000 on-line things one can face would be far more subtle. Suppose one user arrives at a T-shirt channel, in spite of what she has purchased any merchandise, whether or not she is intensively actuated to shop for one thing this point will extremely have an

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

on the $64000 recommendation result. Below this circumstance, the user's temperament, particularly her effect purchase intention would play associate primarily vital role in decisive her judgement to simply accept the things or not. During this paper, we tend to propose a scenario-based approach to check the result of users' purchase

intention on a true recommender system,Tmall.com. Firstly, we tend to statistically analyse the dependence of nineteen representative users' options on their online activity sequence. Secondly, we tend to propose a scenario based approach to severally distinguish users into 2 groups: one with obvious purchase intention, and another while not such motivation.

II. RELATED WORK

Our work is mainly related to three things:

Recommender Systems: Recommender system was defined from the perspective of E-commerce as a tool that helps users search through records of knowledge which is related to user's interest and preference. Recently,various approaches for building recommendation systems have been developed, which can utilize collaborative filtering, content based filtering or hybrid filtering. Collaborative filtering technique is used that recommends items based on users various attributes.
Cross-domain Recommendation: In real world domains could be related to each other by some common information .There are many approaches available for cross domain recommendation, but they are not able to provide better accuracy of high dimensional data and these approaches are suffering from data sparity problem. In this paper we deal with cross-domain recommendation. One of the key techniques for cross domain recommendation is Transfer Learning from the source domain, and further applies it in a target domain.
Social Network Mining: Social network mining is the process of representing, analysing, and extracting actionable patterns and trends from raw social media data.
Our work is built upon these studies, especially the areas of cross-domain and cold-start recommendation**.** Traditional Recommender system only focus on brand-or category-level purchase preference based on a trained classifier ,which cannot be directly applied to our cross site cold-start product recommendation task. In addition their features only include age, gender and Facebook likes, as opposed to a wide range of features explored in our approach. Lastly, they do not consider how to transfer heterogeneous information from social media websites into a form that is ready to use on the ecommerce side, which is the key to address the cross-site cold – start recommendation problem.

III. METHODOLOGY OF PRODUCT RECOMMENDATION

The following are the main attributes are used while designing the product recommender system;

A. Demographic Attributes:

A demographic profile (often shortened as "a demographic")of a user such as sex, age, gender, marital status, education ,career and interests can be used by e-commerce companies to provide better personalised services. We extract users' demographic attributes from their public profiles on Facebook. Demographic attributes have been shown to be very important in marketing, especially in product adoption for consumers.

B. Text Attributes:

Recent studies have revealed that microblogs contain rich commercial intents of users. Also, users' microblogs often reflect their opinions and interests towards certain topics. As such, we expect a potential correlation between text attributes and users' purchase preferences. We perform word segmentation and stop word removal before extracting two types of text attributes below.

Topic distributions: Seroussi et al. proposed to extract topics from user-generated text using the Latent Dirichlet Allocation (LDA) model for recommendation tasks. Follow the same idea, we first aggregate all the microblogs by a

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

user into a document, and then run the standard LDA to obtain the topic distributions for each user. The benefits of topics distributions over keywords are twofold. First, the number of topics is usually set to 50 _ 200 in practice, which largely reduces the number of dimensions to work with. Second, topic models generate condense and meaningful semantic units, which are easier to interpret and understand than keywords.

Word embeddings: Standard topic models assume individual words are exchangeable, which is essentially the same as the bag-of-words model assumption. Word representations or embeddings learned using neural language models help addressing the problem to capture words' contextual semantics. In word embeddings, each dimension represents a latent feature of the word and semantically similar words are close in the latent space. We employ the Skip gram model implemented by the tool word2vec4 to learn distributed representations of words. Finally, we average the word vectors of all the tokens in a user's published document as the user's embedding vector.

C. Network Attributes:

In the online social media space, it is often observed that users connected with each other (e.g., through following links) are likely to share similar interests. As such, we can parse out latent user groups by the users' following patterns assuming that users in the same group share similar purchase preferences.

D. Temporal Attributes:

Temporal activity patterns are also considered since they reflect the living habits and lifestyles of the microblogging users to some extent. As such, there might exist correlations between temporal activities patterns and users' purchase preferences.

Temporal activity distributions: We consider two types of temporal activity distributions, namely daily activity distributions and weekly activity distributions. The daily activity distribution of a user is characterized by a distribution of 24 ratios, and the i-th ratio indicates the average proportion of tweets published within the i-th hour of a day by the user; similarly weekly activity distribution of a user is characterized by a distribution of seven ratios, and the i-th ratio indicates the average proportion of tweets published within the i-th day of a week by the user.

IV. ALGORITHM DETAILS

We have used three different algorithm:

a) LDA(Latent Dirichlet Allocation) algorithm

b)RNIndex algorithm.

c)Sampling.

The main algorithm is LDA algorithm which we have used in our project.It is proposed to extract topics from user-generated text using the Latent Dirichlet Allocation (LDA) model for recommendation tasks. Follow the same idea, we first aggregate all the microblogs by a user into a document,and then run the standard LDA to obtain the topic distributions for each user. The benefits of topicsdistributions over keywords are two fold. First, the number of topics is usually set to 50 ∼ 200 in practice, which largely reduces the number of dimensions to work with. Second, topic models generate condense and meaningful semantic units, which are easier to interpret and understand than keywords.

Alogorithmic steps for LDA:

Step 1: Decide on number of words N the document will have.

Step 2: Choose a topic mixture for the document.

Step 3: Generate each word in the document by:

First picking a topic.

Then using topic to generate the word itself.

For CBOW, the conditional prediction probability is characterized by a softmax function as follows:

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

$$Pr(p_t|\text{context}) = \frac{\exp(\mathbf{v}_{p_t}^{\top} \cdot \mathbf{v}_{context})}{\sum_p \exp(\mathbf{v}_p^{\top} \cdot \mathbf{v}_{context})}.$$

To optimize for computingexponential sum probabilities, hierarchical softmax and negative sampling techniques are commonly used to speed up the training process.

Heterogenous Representation Mapping Using Gradient Boosting Regression Trees

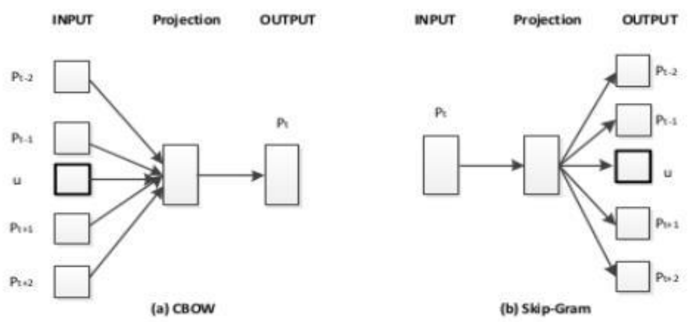


Fig. 1. Two architectures to learn both product and user embeddings. Here u denote a user ID. The major difference between para2vec and word2vec lies in the incorporation of user ID as additional context.

V. MATHEMATICAL MODEL

Let S Let S is the Whole System Consist of

S= {I, P, O}
I = Input.
I = {U, Q, D}
U = User
U = {u1,u2,….,un}
Q = Query Entered by user

Q = {q1, q2, q3,…,qn}

D = Dataset
P = Process:
Step1: Admin will upload the product in E-commerce site.

Step2: That uploaded product will be seen on Social sites where user can view, share and give comments on that product. User can send and receive friend request.
Step3: All the reviews should be seen in E-commerce site when user login to E- commerce site.

O=Ouput: User will get recommendation regarding of that product on ecommerce website.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

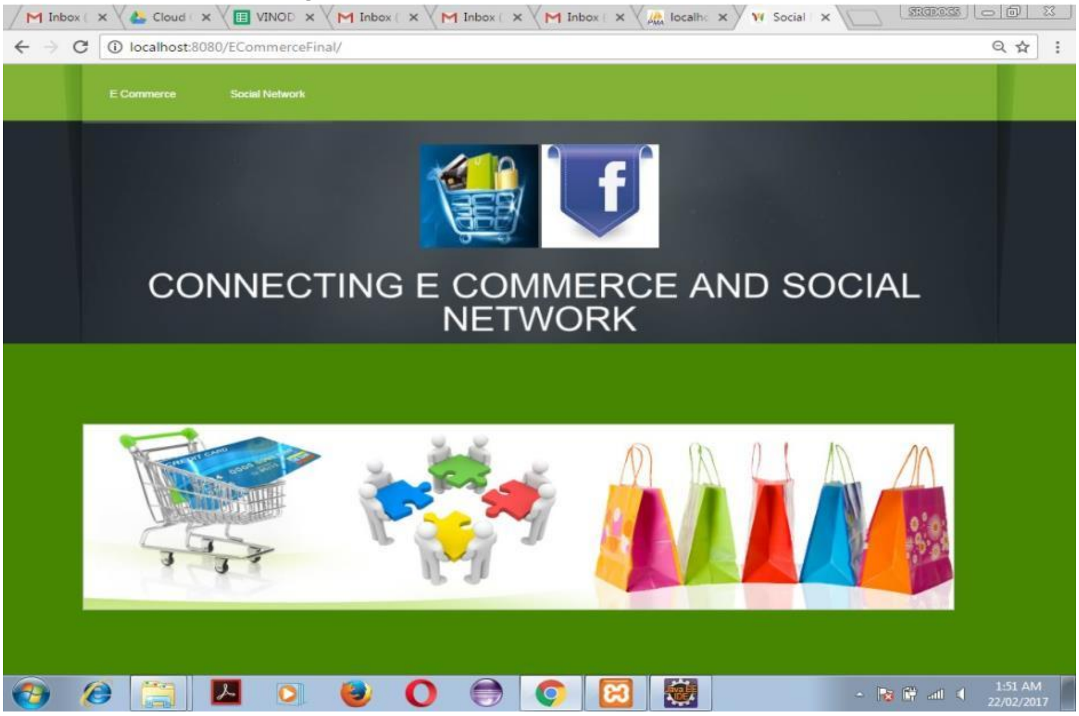
(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

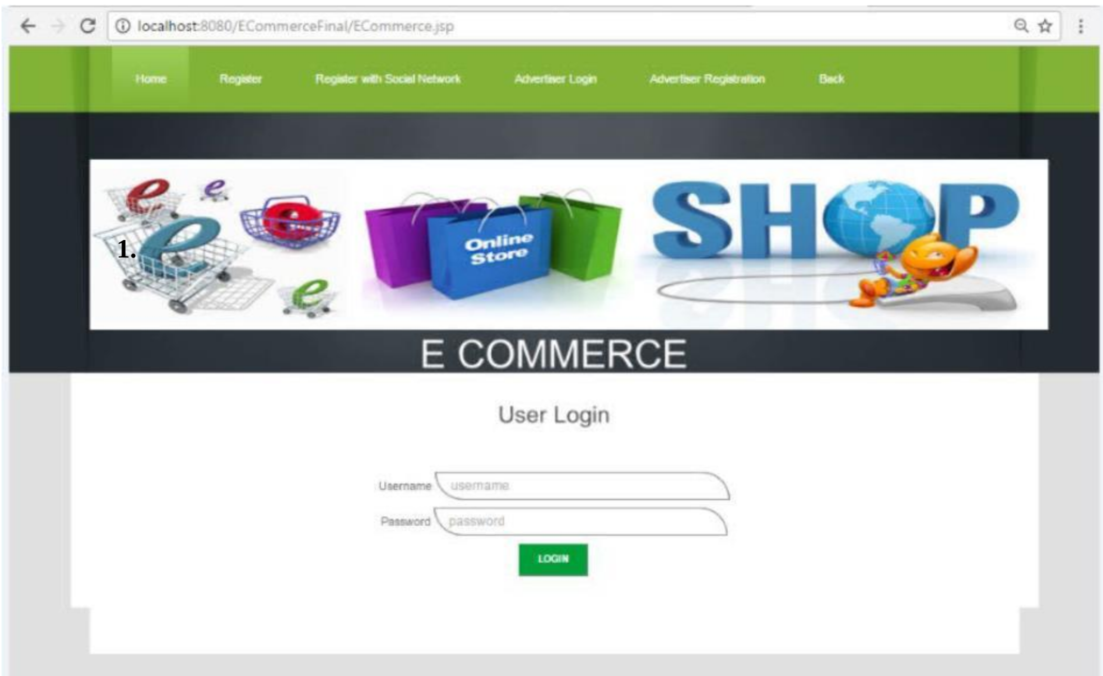
Vol. 5, Issue 3, March 2017

VI. RESULT AND DISCUSSION

1. Home Page:



2. User Login Page:



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

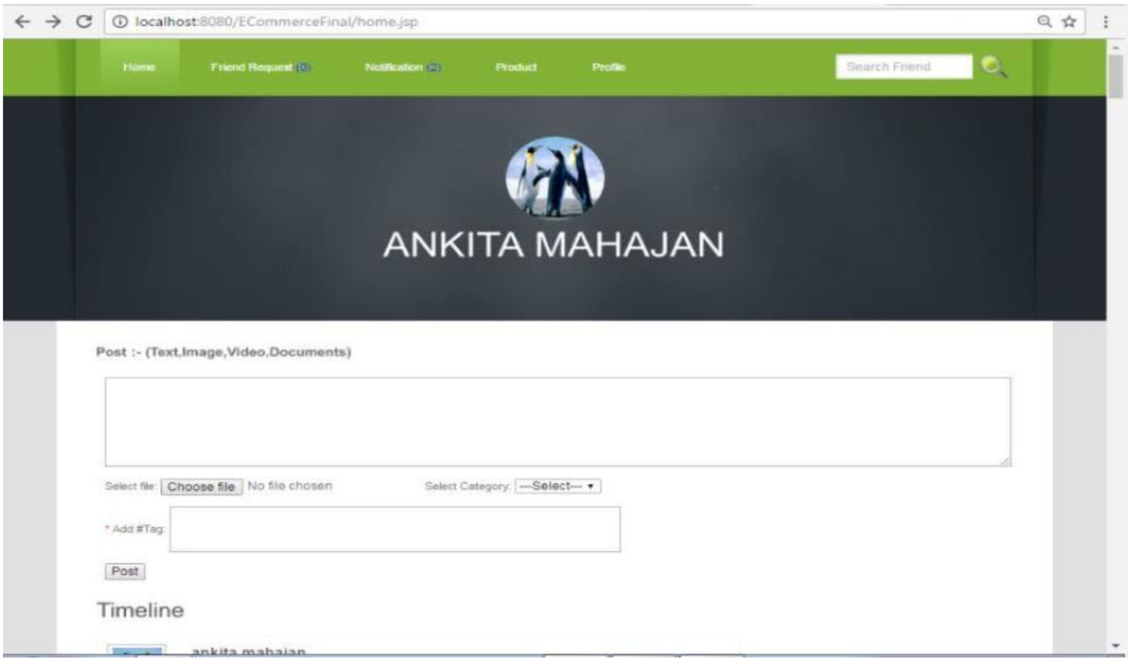
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

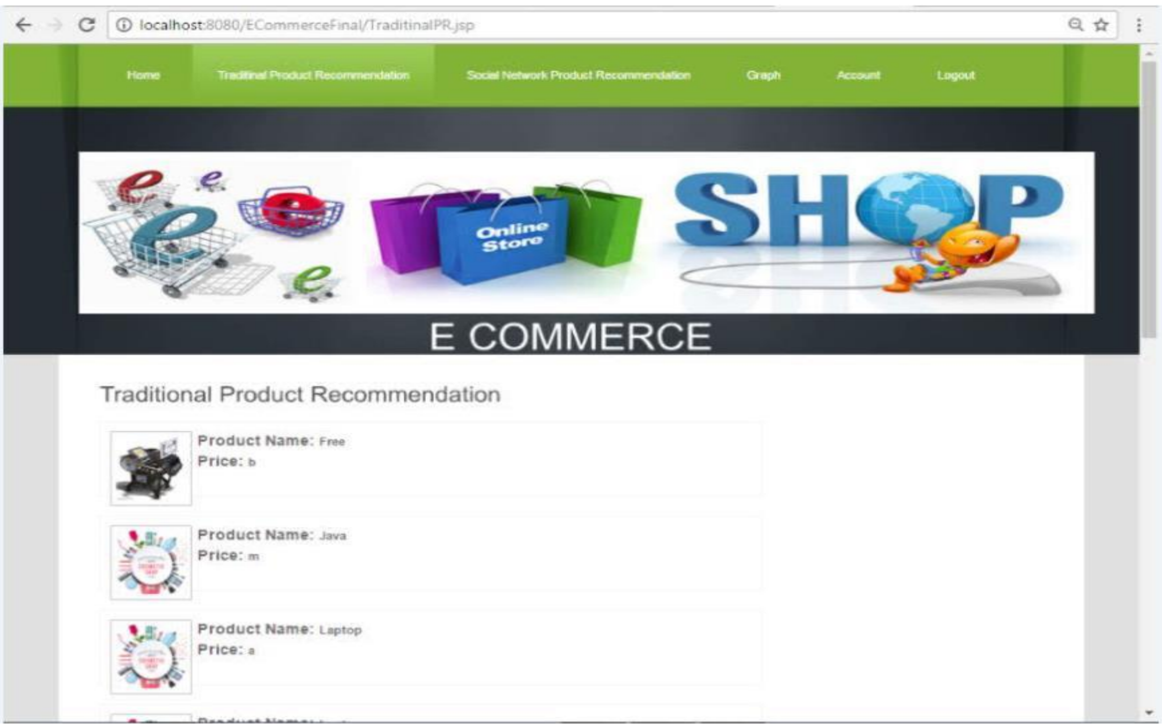
Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

3. User Home Page:



4. E-Commerce Traditional Product Recommendation Page:



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

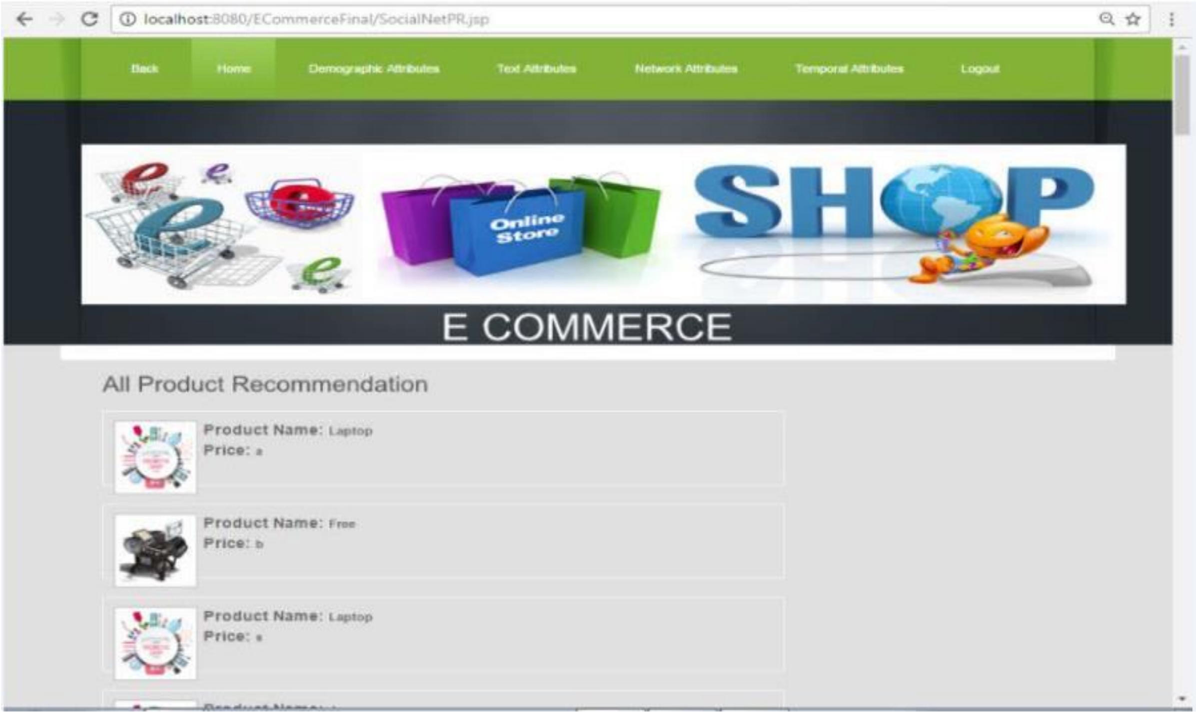
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

5. Social Network Recommendation On E-Commerce:



VII. CONCLUSION AND FUTURE WORK

Recommender systems are turning out to be a useful tool that will provide suggestion to user according to their requirement. Filtering is used to improved recommendation accuracy but each filtering technique associated with some disadvantages with it to overcome this drawbacks here we proposed a recommender system using temporal difference with the help of user feedback (critique) and reinforcement learning we will improve the accuracy of recommendation process. To improve the quality of recommender systems anticipations future research will concentrate on progressing the existing methods and algorithms. Novel lines of research will be formulated for following fields, such as on: (1) The existing recommendation methods that uses different types of available information will be combine in good order, (2) For recommender systems processes enable security and privacy, (3) Flexible frameworks are design for machine-controlled analysis of heterogeneous data.

ACKNOWLEDGMENT

For everything we get, the credit goes to all those who had helped us to complete this project successfully. We are thankful to “Prof. P.B.Sahane” for guidance of this paper. WE would also like to thanks, all faculty members of “PKTC College of Engineering”.

REFRENCES

[1] J. Wang, W. X. Zhao, Y. He, and X. Li, Leveraging product adopter information from online review for product recommendation, in ICWSM, 2015.

[2] W. X. Zhao, Y. Guo, Y. He, H. Jiang, Y. Wu, and X. Li, We know what you want to buy: a demographic-based system for product recommendation on microblogs, in SIGKDD, 2014.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

[3] Q. V. Le and T. Mikolov, Distributed representations of sentences and documents, CoRR, vol.abs/1405.4053, 2014.

[4] J. Lin, K. Sugiyama, M. Kan, and T. Chua, Addressing coldstart in app recommendation: latent user models constructed from twitter followers, in SIGIR, 2013.

[5] J. Wang and Y. Zhang, Opportunity model for e-commerce recommendation: Right product; right time, in SIGIR, 2013.

[6] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, Distributed representations of words and phrases and their compositionality, in NIPS, 2013.

[7] Y. Seroussi, F. Bohnert, and I. Zukerman, Personalised rating prediction for new users using latent factor models, in ACMHH, 2011.

[8] M. Giering, Retail sales prediction and item recommendations using customer demographics at store level, SIGKDDExplor. Newsl., vol. 10, no. 2, Dec. 2008.

[9] G. Linden, B. Smith, and J. York, Amazon.com recommendations: Item-to-item collaborative filtering, IEEE InternetComputing, vol. 7, no. 1, Jan. 2003.

[10] V. A. Zeithaml, The new demographics and market fragmentation, Journal of Marketing, vol. 49, pp.6475, 1985.