ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.7

Cross-Site Cold-Start Product Recommendation for Social Media and E-Commerce Websites

Mayuri Yande[1], Manoj Wakchaure[2]
ME Student[1], Assistant Professor[2]
Department of Computer Engineering
Amrutvahini College of Engineering, Sangamner, India

Abstract:
In recent years the gap between social media and e-commerce reduced tremendously. Many e-commerce websites used social login in status or information for their users' login verification and allow new users to sign in with their existing social login log in information. Both Facebook and Twitter have introduced a new feature last year that allow users to buy products directly from their websites by clicking a "buy" button to purchase items in adverts or other posts. With the new trend of conducting e-commerce activities on social networking sites, it is important to leverage knowledge extracted from social networking sites for the development of product recommender systems. This paper presents a system for predicting user's purchase behaviors on e-commerce websites from the user's social media profile.

Keywords: E-commerce, Matrix Factorization, Microblogs, Product recommender, Recurrent Neural Network, Social media

I. INTRODUCTION

The Internet has changed the way we communicate. New concepts such as e-commerce, Social media, Social marketing, etc. have become more popular and play an important role in linking the information of social media to reach the customer. Recent years have seen increasing boundaries between e-commerce and social networking sites. Some e-commerce websites also support the mechanism of social login where new users can sign in. In this paper, we study the problem which is named as cross site cold start product recommendation. In this problem we recommend products to the customers at social networking sites from an e-commerce website. Recommendation systems have become a core component in today's online business world. The recommendation systems are beneficial in effectively delivering the right items to the right people[17]. A cold-start problem i.e. how to provide recommendations to new users? Is the key challenge for building an effective recommender system [1], [2], [3] .As online shopping becomes popular, e-commerce recommendation is an increasingly important tool. Recommender systems suggest items of interest to users based on their explicit and implicit preferences. In our problem only the users social information is available and we have to transform it for the product recommendation. To address this challenge, we propose to use the users who have social networking accounts and have made purchases on e-commerce websites (i.e. linked users across both sites) as a bridge to transform user's social networking information into the latent features for product recommendation.

II. RELATED WORK

There are three lines of research which relate to proposed system. This is listed below

1) **Recommender system**:-Which make use of the information from users' public profiles and topics extracted from user-generated content into a matrix factorization model for new users' rating prediction.
2) **Cross-domain recommendation:** Collective matrix factorization to estimate the relations of multiple entities by factorizing several matrices simultaneously while sharing parameters in the latent space.
3) **Social network mining:** to route products from e-commerce companies to micro blogging users. Our work is also related to studies on automatic user profiling and cross-site linkage inference.

We highlight related important approaches:
1) Traditional collaborative recommender,
2) Trust-enhanced recommender, and
3) Reviews-based recommender.

First, the traditional collaborative filtering approaches can be either memory-based or model-based. These methods are based on the rating history from users. In the memory-based methods, similarity computation is a primary element. They use a heuristic utility of similarity between users' vectors such as Pearson Correlation Coefficient (PCC) or cosine similarity measure (VCC). On the other hand, the model-based methods employ machine learning models to predict product ratings. For example, Sarwar et al. [4], [18] implemented clustering algorithms to identify groups of customers who rated similar products and these clusters can be seen as likeminded neighbors. Since k clusters are created, recommendation prediction can be computed by averaging the ratings in that cluster.

III. PROBLEM DEFINATION

We deal with the problem of product recommendation of social media users who have unknown history on an e-commerce websites i.e. cold-start situation.

- Build a cold-start recommender system, by providing high-level recommendations to social media users who sign in for the first time to an e-commerce websites.
- Improve existing product recommendation engines that can guide the recommender system to find domains of interest for the user.
- Provide e-commerce companies with tools for targeted social media companies.

IV. SYSTEM ARICHITECTURE

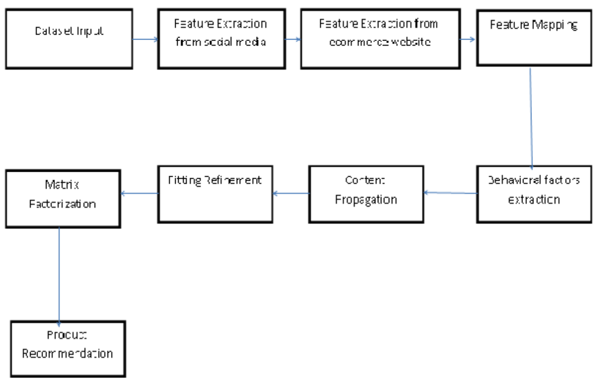

Figure.1.Block diagram of proposed system

System architecture describes the overview and exact flow of working. In this system, first we collect the dataset as an input. The dataset may contain the user's information and list of products. Feature extraction process is done for both the sites i.e. social networking sites and e-commerce websites. These extracted features are mapped for product recommendation using content propagation and matrix factorization algorithm. The flow of the system is sequencial. Matrix factorization technique is used for the representation of the user-item rating matrix.

V. MICROBLOGGING SERVICES

Microblogging Feature selection Methods are –

- Demographic Attributes
- Text Attributes
- Network Attributes
- Temporal Attributes

Table I. Categarisation of the Microblogging Features

Categories	Features
Demographic Attributes	Gender, age, marital status, education, career, etc.
Text Attributes	Topic distribution, Word embedding
Network Attributes	Latent group preferences
Temporal Attributes	Daily activity distribution, weekly activity distribution

VI. PERFORMANCE ANALYSIS

We use the users in Ddense as the training data for both user embedding fitting and matrix factorization learning, and consider the users in Dsparseas the test data for product recommendation. Since the users in Dsparse have fewer than five purchases, we only report the performance of Recall@k butnot Precision@k. We also use MAP, MRR and AUC as evaluation metrics. We can observe that our proposed method ColdE is consistently better than all the baselines, which indicates that the effectiveness of recommendation for long-tail users.

Table.2. Performance Comparision On Cold Start Product Recommendation

Methods	R@10	AUC
Pop	0.120	0.684
Pop++	0.120	0.684
MFUA	0.415	0.715
FMUI	0.419	0.718
$Cold_E$	0.458	0.757

Methods of comparision:

- **Popularity(Pop)**: products are ranked by their historical sale volumes.
- **Popularity with semantic similarity(pop++)**: the ranking score is a product of popularity score and the cosine similarity.
- **MF with user attribute(MFUA):** matrix factorization algorithm is used to incorporate the user attributes.
- **FM without user interactions(FMUI):** this feature is equivalent to SVD feature.
- **Cold $_E$:** Our proposed approach, here we use the fitted user embedding features and product embedding features.

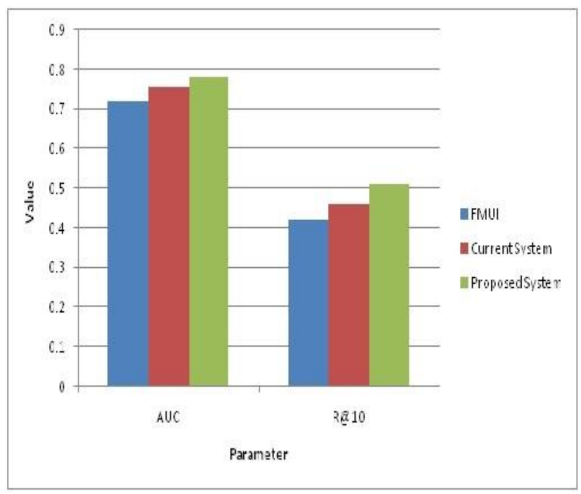

Figure.2. Comparision of current and proposed system

VII. OUR CONTRIBUTION

Micro blogging user adopts some content propagated to her, we can attribute that to three behavioral factors, namely, topic virality, user virality and user susceptibility. Topic virality measures the degree to which a topic attracts propagations by users. User virality and susceptibility refer to the ability of a user to propagate content to other users, and the propensity of a user adopting content propagated to her, respectively.

- Formulate the problem of recommending products from an e-commerce website to social networking users in "cold-start" situations.
- apply the recurrent neural networks for learning correlated feature representations for both users and products.

- Propose a modified gradient boosting trees method to transform users' microblogging attributes to latent feature representation which are used for product recommendation.
- a feature-based matrix factorization approach is used by incorporating user and product features for cold-start product recommendation.

VIII. CONCLUSION

In our Approach, we have studied a novel problem, cross-site cold-start product recommendation, i.e., recommending products from e-commerce websites to microblogging users. The mapped user features can be effectively incorporated into a feature-based matrix factorisation approach .Our study will have profound impact on both research and industry communities. Our main idea is that on the e-commerce websites, users and products can be represented in the same latent feature space through feature learning with the recurrent neural networks. Using a set of linked users across both e-commerce websites and social networking sites as a bridge, we can learn feature mapping functions using a modified gradient boosting trees method, which maps users' attributes extracted from social networking sites onto feature representations learned from e-commerce websites. The results show that our proposed framework is indeed effective in addressing the cross-site cold-start product recommendation problem.

IX. REFERENCES

[1]. Wayne Xin , Sui Li, Yulan He, Edward Y.Chang,Ji-Rong Wen, Xiaoming Li, "Connecting Social media to E-commerce: Cold-Start Product Recommendation using Microblogging Information", IEEE Transaction on Knowledge and Data Engineering, VOL.2,NO.X,XXX 2015.

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

[3].M. Giering, "Retail sales prediction and item recommendations using customer demographics at store level," *SIGKDD Explor. Newsl.*, vol. 10, no. 2, Dec. 2008.

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

[5].V. A. Zeithaml, "The new demographics and market fragmentation," *Journal of Marketing*, vol. 49, pp. 64–75, 1985.

[6]. Y. Seroussi, F. Bohnert, and I. Zukerman, "Personalised rating prediction for new users using latent factor models," in *ACM HH*, 2011.

[7]. 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.

[8] .T. Chen, H. Li, Q. Yang, and Y. Yu, "General functional matrix factorization using gradient boosting," in *ICML*, 2013

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

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

[11]. T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *CoRR*, vol.abs/1301.3781, 2013.

[12]. Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[13]. J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, vol. 29, pp. 1189–1232, 2000.

[14]. L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Monterey, CA: Wadsworth and Brooks,1984.

[15]. L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, Oct.2001.

[16]. K. Zhou, S. Yang, and H. Zha, "Functional matrix factorizations for cold-start recommendation," in *SIGIR*, 2011.

[17]. T. Chen, H. Li, Q. Yang, and Y. Yu, "General functional matrix factorization using gradient boosting," in *ICML*, 2013.

[18] .T. Chen, W. Zhang, Q. Lu, K. Chen, Z. Zheng, and Y. Yu, "SVD Feature: A toolkit for feature-based collaborative filtering," *Journal of Machine Learning Research*, vol. 13, 2012.

[19].S. Rendle, "Factorization machines with libfm," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 3, May 2012.