ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol.3, Special Issue.24, March 2017

Enhanced Architecture Design For Social Media E-commerce And News Using Advanced Microblogging Information

A.Briska, K.S.Dhivya, S.Najima banu
Final CSE
T. Poornima M.E., AP/CSE
Maharaja Engineering College, Avinashi.

Abstract— The boundaries between e-commerce and social networking have become increasingly blurred. Many e-commerce websites support the mechanism of social login where users can sign on the websites using their social network identities such as their Facebook or Twitter accounts. Users can also post their newly purchased products on microblogs with links to the e-commerce product web pages. we propose a novel solution for cross-site cold-start product recommendation, which aims to recommend products from e-commerce websites to users at social networking sites in "cold-start" situations, a problem which has rarely been explored before. A major challenge is how to leverage knowledge extracted from social networking sites for cross-site cold-start product recommendation. We propose to use the linked users across social networking sites and e-commerce websites (users who have social networking accounts and have made purchases on e-commerce websites) as a bridge to map users' social networking features to another feature representation for product recommendation. In specific, we propose learning both users' and products' feature representations (called user embedded and product embedding, respectively) from data collected from e-commerce websites using recurrent neural networks and then apply a modified gradient boosting trees method to transform users' social networking features into user embedding. We then develop a feature-based matrix factorization approach which can leverage the learnt user embedding for cold-start product recommendation. Experimental results on a large dataset constructed from the largest Chinese microblogging service SINA WEIBO and the largest Chinese B2C e-commerce website JINGDONG have shown the effectiveness of our proposed framework.

1. INTRODUCTION

IN recent years, the boundaries between ecommerce and social networking have become increasingly blurred. Ecommerce websites such as eBay features many of the characteristics of social networks, including real-time status updates and interactions between its buyers and sellers. Some ecommerce websites also support the mechanism of social login, which allows new users to sign in with their existing login information from social networking services such as Facebook, Twitter or Google+. 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. In China, the ecommerce

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol.3, Special Issue.24, March 2017

company ALIBABA has made a strategic investment in SINA WEIBO1 where ALIBABA product adverts can be directly delivered to SINA WEIBO users. 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. In this paper, we study an interesting problem of recommending products from ecommerce websites to users at social networking sites who do not have historical purchase records, i.e., in "cold-start" situations. We called this problem cross-site cold-start product recommendation. Although online product recommendation has been extensively studied before most studies only focus on constructing solutions within certain e-commerce websites and mainly utilise users historical transaction records. To the best of our knowledge, cross-site cold-start product recommendation has been rarely studied before. In our problem setting here, only the users social net- working information is available and it is a challenging task to transform the social networking information into latent user features which can be effectively used for product recommendation. To address this challenge, we propose to use the linked users across social networking sites and ecommerce websites (users who have social networking accounts and have made purchases on e-commerce websites) as abridge to map users social networking features to latent features for product recommendation. In specific, we pro- pose learning both users and products' feature representations (called user embedding and product embedding, respectively) from data collected from e-commerce websites using recurrent neural networks and then apply a modified gradient boosting trees method to transform users' social networking features into user embedding. We then develop a feature-based matrix factorization approach which can leverage the learnt user embedding for cold- start product recommendation. We built our dataset from the largest Chinese microblogging service SINA WEIBO2 and the largest Chinese B2C e-commerce website JING DONG, containing a total of 20,638 linked users. The experimental results on the data shown the feasibility and the effectiveness of propose framework. Our major contributions are summarised below:

We formulate a novel problem of recommending products from an e-commerce website to social net- working users in "cold-start" situations. To the best of our knowledge, it has been rarely studied before. We propose to apply the recurrent neural networks for learning correlated feature representations for both users and products from data collected from an e-commerce website. We propose a modified gradient boosting trees method to transform users microblogging attributes to latent feature representation which can be easily incorporated for product recommendation. We propose and instant feature- based matrix factorization approach by incorporating user and product features for cold start product recommendation.

2. PROBLEM FORMULATION

Given an e-commerce website, let U denote a set of its users, P a set of products and R purchase record matrix, each entry of which is a binary value indicating whether u has purchased product p. Each user u 2Uis associated with a set of purchased products with the purchase time- stamps. Furthermore, a small subset of users in U can be linked to their microblogging accounts (or other social net- work accounts), denoted U .As such, each user u 2UL is also associated with their respective microblogging attribute information. Let A denote the set of microblogging features, and each

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol.3, Special Issue.24, March 2017

microblogging user has k-dimensional micro- blogging feature vector au, in which each entry is the attribute value for the in microblogging attribute feature.

With the notations introduced above, we define our recommendation problem as follows. We consider a cross-site cold-start scenario: a microblogging user u0 =2Uis new to the e-commerce website, who has no historical purchase records. It is easy to see u0 =2UL, too, since have we aim to generate a personalised ranking of recommended products for u0 based on her Microblogging attributes.

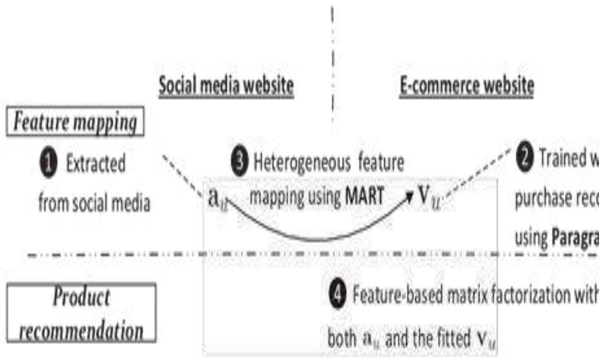


Fig. 1. The workflow diagram for our presented solution.

Services cannot usually be used directly for product recommendation on e-commerce websites. Therefore, one major challenge is how to transform users' microblogging attribute information au0 into another feature representation which can be used more effectively for product recommendation. Here, we call the original or microblogging feature representation and vu0 the (heterogeneous) transformed feature representation, respectively. Next, we will study how to extract microblogging features and transform them into a distributed feature representation before presenting a feature-based matrix factorization approach, which incorporates the learned distributed feature representations for product recommendation. The entire workflow of our solution is shown in Fig. 1 which consists of four major steps splitting into feature mapping and product recommendation, which will be discussed in Sections 3 and 4 respectively

3. EXTRACTING AND REPRESENTING MICROBLOGGING ATTRIBUTES

Our solution to microblogging feature learning consists of three steps: 1.Prepare a list of potentially useful microblogging attributes and construct the microblogging feature vector au for each linked user u 2UL; 2.Generate distributed feature representations fvugu2U using the information from all the users U on the e-commerce website through deep learning; Learn the mapping function, which transforms the microblogging attribute information au to the distributed feature representations vu in the second step. It utilises the feature representation pairs of all the linked users 2UL as training data.

Microblogging Feature Selection

In this section, we study how to extract rich user information from microblogs to construct au for a microblogging user. We consider three groups of attributes.

Demographic Attributes

A demographic profile (often short end as a demographic") of a user such as sex, age and education can be used by e-commerce companies to provide better personalised services. We extract users' demographic attributes from their public profiles on SINA WEIBO. Demographic attributes have been shown to be very important in marketing, especially in product adoption for consumers Following our previous study we identify six major demographic attributes: gen- der, age, marital status, education, career and interests. To quantitatively measure these attributes, we have further discretized them into different bins following our previously proposed method.

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 Chinese word segmentation and stop word removal before extracting two types of text attributes below. Topic distributions. The proposed to extract topics from user-generated text using the Latent 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 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 embedding. Standard topic models assume individual words are exchangeable, which is essentially the same as the bag-of-words model assumption. Word representations or embedding learned using neural language models help addressing the problem.

Traditional bag-of- word approaches which fail to capture words' contextual semantics In word embedding, 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.

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. Latent group preference. Since it is infeasible to consider all users on WEIBO and only keeping the top users with the most followers would potentially miss interesting information, we propose to use topic models to learn latent groups of follow- We treat a following user as a token and aggregate all the followings of a user as individual document .In this way, we can extract latent user groups sharing similar interests (called "following topics"), and we represent each user as a preference distribution over these latent groups.

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 characterised by a distribution of 24 ratios, and the ratio indicates the average proportion of tweets published within the of a day by the user; similarly weekly activity distribution of a user is characterised by a distribution of seven ratios, and the ratio indicates the

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol.3, Special Issue.24, March 2017

average proportion of tweets published within the day of a week by the user.

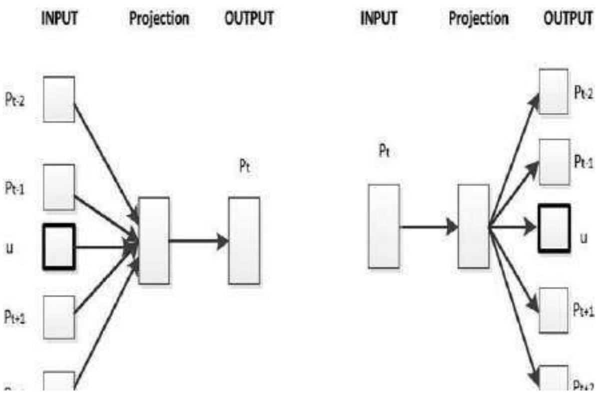
3.2 Distributed Representation Learning with Recurrent Neutral Networks

we have discussed how to construct the microblogging feature vector au for a user u. However, it is not straightforward to establish connections between au and products. Intuitively, users and products should be represented in the same feature space so that a user is closer to the products that she has purchased compared to those she has not. Inspired by the recently proposed methods in learning word embedding using recurrent neutral networks we propose to learn user embedding or distributed representation of user vu in a similar way.

Learning Product Embedding

Before presenting how to learn user embedding, we first discuss how to learn product embedding. The neural net- work methods, word2vec, proposed in for word embedding learning can be used to model various types of sequential data. The core idea can be summarised as follows. Given a set of symbol sequences, a fixed-length vector representation for each symbol can be learned in a latent space by exploiting the context information among symbols, in which “similar” symbols will be mapped to nearby positions. If we treat each product ID as a word token, and convert the historical purchase records of a user into a time stamped sequence, we can then use the same methods to learn product embedding. Unlike matrix factorization, the order of historical purchases from a user can be naturally captured. We consider two simple recurrent neutral architectures proposed in to train product embedding, namely, the Continuous Bag-Of-Words model (CBOW) and the Skip-gram model. The major difference between these two architectures lies in the direction of prediction: CBOW predicts the current product using the surrounding context, i.e. while Skip-gram predicts the context with the current product, In our experiments, the context is defined as a window of size 4 surrounding a target product which contains two products purchased before and two after pt. More formally, each product is model as a unique latent embedding vector and the associated context vector is obtained to average the vectors of the context information as v context.

For CBOW, the conditional prediction probability is characterized by a soft max function. To optimize for computing exponential sum probabilities, hierarchical soft max and negative sampling techniques are commonly used to speed up the training process. At each training iteration, we sample a target product together with their context window, and then update the parameters with Stochastic Gradient Descent (SGD) using the gradients derived by back Learning for Skip-gram is done in a similar way, which is omitted here



ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol.3, Special Issue.24, March 2017

Fig. 2. Two architectures to learn both product and user embedding. Here u denote a user ID. The major difference between para2vec and word2vec lies in the incorporation of user ID as additional context.

Learning User Embedding

Given product embedding, if we can learn user embedding in a similar way, then we can explore the correlated representations of a user and products for product recommendation. We borrow the idea from the recently proposed Paragraph Vector method [9], which learns feature representations from variable-length pieces of texts, including sentences, paragraphs, and documents. We implement a simplified version of para2vec at the sentence level as follows. The purchase history of a user can be considered as a "sentence" consisting of a sequence of product IDs as word tokens. A user ID is placed at the beginning of each sentence, and both user IDs and product IDs are treated as word tokens in a vocabulary in the learning pro- cess. During training, for each sentence, the sliding context window will always include the first word (i.e., user ID) in the sentence. In this way, a user ID is essentially always associated with a set of her purchase records (a context window of four products at a time). We can then use the same learning procedure in word2vector for the estimation of and we present an illustrative example of these two architectures in Fig. 2. After learning,

We separate user embedding from product embedding and use vu and to denote the learnt K-dimensional embedding for user u and product p respectively. The rationales of applying para2vec to model purchase data can be explained below. First, the user embedding representation for each user ID reflects the users' personalized purchase preference; second, the surrounding context, i.e., product purchases, is used to capture the shared purchase patterns among users. Compared to the traditional matrix factorization the (window-based) sequential context is additionally model in addition to user preference, which is expected to potentially yield better recommendation results.

Heterogenous Representation Mapping technique

Using Gradient Boosting Regression Trees We have presented how to construct a microblogging feature vector au from a microblogging site and learn a distributed representation vu from an e-commerce website respectively. In the cross-site cold-start product recommendation problem we considered in this paper (i.e., make a product recommendation to a user u who has never purchased any products from an e-commerce website), we can only obtain the microblogging feature vector au for user u. The key idea is to use a small number of linked users across sites as a bridge to learn a function which maps the original feature representation au to the distributed representation vu. Specifically, we can construct a training set consisting of feature vector pairs, and cast the feature map- ping problem as a supervised regression task: the input is a microblogging feature vector au and the output is a distributed feature vector vu. Assume that vu contains K dimensions, we need to learn a set of K function, and the function takes the original feature vector of a user u as the input and returns the corresponding transformed feature value We extend the Multiple Additive Regression Tree (MART) [13] method to learn feature map- ping functions since it is powerful to capture higher-order transformation relationship between input

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol.3, Special Issue.24, March 2017

and output.

A Brief Introduction of MART

Gradient boosting algorithms aim to produce an ensemble of weak models that together form a strong model in a stage-wise process. Typically, a weak model is a J-terminal node Classification and Regression Tree (CART) and the resulting gradient boosting algorithm is called Multiple Additive Regression Tree. An input feature vector x 2 Rd is mapped to a score FðxÞ2R. The final model is built in a stage-wise process by per- forming gradient descent in the function space. At the boosting.

Completeness-Based Feature Sampling

An issue about the gradient boosting algorithm is that it tends to over fit the training data. It has been previously shown that the incorporation of randomized feature sampling improves the tree based ensemble methods in Random
Forest Inspired by the idea, we propose to use an attribute-level importance sampling method where each attribute is assigned with an importance score and at each node split in building the MART trees, we only sample a fraction of attributes (empirically set to 2 3) based on each attribute's importance score instead of enumerating all the attributes. Once an attribute is sampled, its corresponding attribute value features will be selected subsequently. The importance score of each attribute is set to the proportion of the attribute values that can be extracted from the users' public profiles on SINA WEIBO. Another benefit of complete- ness-based sampling is that attributes with a larger proportion of missing values will be more likely to be pushed to the leaf nodes, which alleviates the missing value problem in regression trees.

4. EVALUATION FOR COLD START RECOMMENDATION

We aim to recommend products to microblog users without the knowledge of their historical purchase records. Construction of the Evaluation Set the evaluation set splits users into training set and test set. For the training set we sample negative products with a ratio of1 for each user, we have the same number of negative and positive products. For the test set, we randomly sample negative products with a ratio of 1:50 for each user, i.e., each positive product would involve 50 negative products. All negative products are sampled from the same product category as the corresponding positive one. For example, for "iPhone 6", we can sample "Samsung Galaxy S5" from the "Mobile Phones" category as a negative product. Given a user, we can generate a list of candidate products consisting of both positive and negative products. On average, a user has about 52 positive products and 2,600 negative products in our experimental dataset, which is indeed a challenging task. Similar to the evaluation scenario in Information Retrieval, we would like to examine the performance that a system ranks positive product some negative products.

4.1. Methods to Compare We consider the following methods for performance comparison:

Popularity (Pop): products are ranked by their historical sale volumes. Popularity with Semantic Similarity (Pop++): the ranking score is a combination of two scores: (1) the popularity score S1; (2) the cosine similarity S2 between product description and user text information, including profile, tweets and tags. The Embedding Similarities (ES): Similarity scores vu and a list of between a user embedding product embedding are used to rank products. With user attributes (MFUA): User attributes (including user profile and topic distributions) are incorporated into the basic matrix factorisation algorithm for product rating prediction, we also use the pairwise loss

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol.3, Special Issue.24, March 2017

function to train the model. FM without User Interactions (FMUI): applied the Factorization Machines word embedding are used to capture the semantic characteristics of user-generated text, they have different focuses. Topic distributions are more suitable to extract topical themes from text based on word co-occurrence patterns (essentially taking the whole document as the context window) while word embed- dings are more suitable to capture the semantics between words from local context windows, usually comprising three words before and after the target word. Hence, we keep both types of text features in our approach. It is worth noting that our method is a tree-based approach, which can effectively handle information redundancy, i.e., if a feature contains redundant information given the tree that is being constructed, it will be pushed to a lower rank during attribute selection. 8. In our dataset, the completeness proportion of demographic attributes are as follows: Gender (100 percent), Interests (65.7 percent), Age (36.7 percent), Education (26.3 percent), Career (12.9 percent) and Marital status (4.6 percent); while for text and network attributes, the proportion of completeness.

About 99.1 percent, i.e., most users have published tweets and followed some other users.

Recommendation in KDDC up 2012. It has been found that similar performance was obtained with or without the interactions of user features. FM without user feature interactions is equivalent to SVD Feature. Were implement this method in the SVD Feature framework with our extracted micro- blogging features. : Our proposed approach which uses the fitted user embedding features and product embedding features. Our proposed approach which uses the microblogging features, the product embedding features and the fitted user embedding features. Especially, we only use demographic attributes here, since they have been shown important to product recommendation. Since the user and product embedding can be learned for all the users and products respectively in the e-commerce website, we can train with all the users in U, not limited to the linked users UL. This variant is called Colden method. We set the regularization coefficient to a 0:004, the iteration number to 50 and the factor number to 32 for all the methods. We use the CBOW architecture to learn the embed- ding vectors based on the purchase records from all the non-linked users and the partial purchase records from linked users in our training set. The number of dimensions of embedding vectors is set to 50. The user embedding features in the test sets for different #training #test settings are set to the values fitted using MART both. We add additional 10,000 randomly selected non-linked users from U into the training set. 5.4 Revisiting the Effectiveness of the Distributed Representations of Users and Products In the previous section, we have shown that the learn product and users embedded are effective to improve the recommendation performance. In this section, we give more in sights in to the effectiveness so the distributed representations.

Insights into Product Embedded First, we take the learnt product embedded to conduct a quantitative similarity analysis in order to find out whether the learned product embedded can discriminate products from different categories or brands. We compute the average similarity score between product pairs from (1) different categories and brands (DCDB); (2) same category but different brands (SCDB); and (3) same category and same brand (SCSB). As it is infeasible to calculate the similarity scores for all possible product pairs in JINGDONG, we sample 10 million product pairs randomly for each type of product pairs for computation. The results are as follows: DCDB ¼ 0:0217, SCDB ¼ 0:2719 and SCSB ¼ 0:4406. The average similarity score of SCDB > indicates the product embedding learned are indeed very different for products under different categories; while SCSB > SCDB indicates the product embedding have a good discriminative power for brands. Insights into User Embedding we take the learnt user embedding to conduct a quantitative similarity analysis in order to find out whether the learned user embedding can identify users with similar purchase history. Given a user u, we build two groups of users, denoted by GA u and GB. GA u contains the top K most similar users (a.k.a. K nearest neighbours) of user u, which are identified by the coefficient in terms of purchase history; GB u contains K randomly selected users. We would like to examine whether the user embedding vectors can discriminate a user in GA u from another one in GB. Given user u together with GA u and GB u we can derive two similarity values A and which are the average similarities with the users in GA u and the users in GB u respectively for user u. We use the cosine function to compute the similarity between two user embedding vectors. K is set to 30 in our experiments. In this way, we can obtain two arrays of similarity values by constructing the paired t-test, the results have shown that the values in A gu2U are significantly larger than those in gu2U at the level of 0.001. The average similarities for Agu2U B gu2U are 0.090 and 0.031 respectively.

5. RELATED WORK

Our work is mainly related to three lines of research: Recommender systems. In recent years, the matrix factor approach has received much research interests. With the increasing volume of web data, many studies focus on incorporating auxiliary information into the matrix factorization approach. Two typical frame- works of such studies are the SVD Feature and Factorization Machine . There has also been a large body of research work focusing specifically on the cold-start recommendation problem. Serous proposed to 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. Zhang et al. propose a semi- supervised ensemble learning algorithm. Schein pro- posed a method by combining content and collaborative data under a single probabilistic framework. Lin et al. addressed the cold-start problem for App recommendation by using the social information from Twitter. Tension et al. Zhou et al. experimented with eliciting new user preferences using decision trees by querying users' responses progressively through an initial interview process. Mosh proposed a method for combining content features such as semantic and emotion information with ratings information for the recommendation task. Chang presented an influence-based diffusion model considering user influence in addition to relevance for matching ads. Liu et al. identified representative users whose linear combinations of tastes are able to approximate other users. Cross-domain recommendation. One of the key techniques for cross-domain recommendation is Transfer Learning and the idea is to learn transferable knowledge from the source domain, and further apply it in a target domain .Singh and Gordon proposed collective matrix factorization to estimate the relations of multiple entities by factorizing several matrices simultaneously while sharing parameters in the latent space. Li attempted to transfer user-item rating patterns from an auxiliary matrix in another domain to the target domain through Codebooks. Hu and Zhao extended transfer learning to triadic factorization and active learning for cross-domain recommendation, respectively. Social network mining. We follow the early commercial mining studies on social networking websites. Holler et al. presented the first work on commercial intent detection in Twitter. Zhao et al. first proposed to route products from e-commerce companies to microblogging users. Our work is also related to studies on automatic user profiling and cross-site linkage inference . Our work is built upon

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol.3, Special Issue.24, March 2017

these studies, especially in the areas of cross-domain and cold-start recommendation. Though sharing some similarities, we are dealing with a very specific task of highly practical value, cold-start product recommendation to microblogging users. To the best of our knowledge, it has not been studied on data set before. Relevant studies are from by connecting users across eBay and Facebook. However, they 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 gender, age 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 in to a form that is ready for use on the e-commerce side, which is the key to address the cross-sitecold-startrecommendationproblem.12.
All the improvement is significant at the confidence level of 0.01.

6.CONCLUSIONS

We have studied a novel problem, cross-site cold-start product recommendation, i.e., recommending products from e-commerce websites to microblogging users without historical purchase records. 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 mapped user features can be effectively incorporated into a feature-based matrix factorisation approach for cold-start product recommendation. We have constructed a large dataset from WEIBO and JINGDONG. The results show that our proposed framework is indeed effective in addressing the cross-site cold-start product recommendation problem. We believe that our study will have profound impact on both research and industry communities. Currently, only a simple neutral network architecture has been employed for user and product embedding learning. In the future, more advanced deep learning models such as ConvolutionalNeuralNetworks13 can be explored for feature learning. We will also consider improving the current feature mapping method through ideas in transferring learning.

REFERENCES

[1] J. Wang and Y. Zhang, "Opportunity model for E-commerce recommendation: Right product; right time," in Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2013, pp. 303–312.

[2] M. York, "Retail sales prediction and item recommendations using customer demographics at store level," SIGKDD Explore News., vol. 10, no. 2, pp. 84-89, Dec. 2008.

[3] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering," IEEE Internet Compute, vol. 7, no. 1, pp. 76–80, Jan / Feb. 2003.

[4] V. A. Linden "The new demographics and market fragmentation," J. Marketing, vol. 49, pp. 64–75, 1985.

[5] W. X. Zhao, Y. Wang, 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 Proc. 20th ACM SIGKDD Int. Conf. Know. Discovery Data Mining, 2014, pp. 1935–1944.