International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Social Recommendation for Interactive Online System

Thenmalar V[1], Tamilselvi R[2], Sandhiya S[3], Dhivya Shree M[4]

[1, 2, 3]UG Scholar, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore

[4]Assistant Professor, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore

Abstract: ***The research communities of information retrieval, machine learning and data mining has attracted lot of attention by a social recommendation system. Online Shopping has become a popular trend as there is no need to go hunting for products in retail markets. Social recommendation plays important role in online system where data arrives sequentially and user preference may change rapidly. To improve the linear optimization of the system using OGRPL-FW this utilizes Frank Wolfe algorithm. This Linearization induces a sublinear rate of convergence and zigzagging behavior. In this paper, we present a new framework in order to avoid these problem a regularization penalty term is added. Using collaborative filtering as well as association rules to find the missing values and predict the items efficiently. It also provides Web Based Application with user-friendly interface. The communication among the customers is provided through online chat facility, wherein they can discuss about the products for reviews. Recommendation system through web is necessary to meet the dynamically changing user preferences.***

Keywords: OGRPL-FW, Regularization Penalty, Association Rules, Collaborative filtering, Chat facility

1. Introduction

A recommendation engine, also known as a recommender system, is software that analyses available data to make suggestions for something that a website user might be interested in, such as a book, a video or a job, among other possibilities. A search-engine is one type of recommendation engine, responding to search queries with pages of results that are (at least theoretically) the best suggestions for websites that satisfy the user's query, based on the search term plus other data, such as location and trending topics. In literature, a variety of social recommendation models are proposed, which can be generally grouped in two categories: matrix factorization based methods and probabilistic model based methods. The methods of both categories are trained from the partially observed user-item matrix and users' social relations. Most social recommendation algorithms are based on batch learning techniques which assume all user ratings are provided in the user-item matrix. Such assumption makes them unsuitable for real-world online recommendation applications. First, the user ratings arrive sequentially in an online application. The batch recommendation algorithm has to be retrained from scratch whenever new ratings are received, making the training process extremely time-consuming. Moreover, if the size of training data is too large, it is difficult for handling all the data in the batch mode. Second, it is common that user preference could drift over time in real world online application, which makes the batch learning processes fail to capture such changes on time. To overcome these difficulties, we develop a novel framework of social recommender system termed Online Graph Regularized User Preference Learning (OGRPL). In the task of online recommendation, the number of user ratings collected at each timestamp is much smaller than the ratings in the offline recommendation, which means all the items have to be recommended in a cold-start manner. Currently, social networking and knowledge sharing sites like Twitter and Douban are popular platforms for users to generate shared opinions for the items like item review and summary. Thus, the content generated by users provides the auxiliary information for the items, which has been widely used to tackle the problem of cold-start item. Thus, the direct learning of user preference may be over-fitting and is therefore not robust. To overcome the over-fitting problem, we formulate the problem of user preference learning with low-rank constraints and learn the low-rank representation of user preference. The OGRPL model recommends the items based on user preference in the online manner. When the recommended items come, users give the rating to the items. The users' ratings are sequentially collected and stored in the system. Then, the OGRPL model updates the user preference based on the newly observed users' ratings and their social relations.

2. Literature Review

Data mining techniques can be implemented at a great rate on existing software and hardware platforms to enhance the value of existing information resources and can be integrated with new products and systems as they are brought online. Due to the rapid growth of information on the Web, especially on the social Web, recommender system has become an indispensable technique for filtering and recommending online information. In order to satisfy Web users ever-increasing information needs, traditional recommendation techniques have been widely adopted by various products in industrial companies, including but not limited to Amazon, Netflix, Apple iTunes, Yahoo! News, etc. Traditional recommendation techniques normally only take into account the user-item rating matrix for computing recommendations. Recently, based on the intuition that users' social network information can be utilized to improve recommendation qualities, the research of social recommender systems becomes popular. Several social recommendation approaches have been proposed in the literature. These methods suggest that the explicit social information is very helpful in improving the traditional methods, especially when the user-item rating matrix is sparse.

3. The Problem of Online Social Recommendation

In this section, we first introduce some notations used in the subsequent discussion, which are the rating matrix R, the feature content matrix of items X, the similarity matrix of users' social relations S and the target preference matrix W. We next present the problem of online social recommendation from the viewpoint of online graph regularized user preference learning. We then learn the user preference using both the collaborative user-item relationship and item content features as an unified learning process. We represent the item content feature in recommender systems using bag-of-words model. We take the item review by users in the recommender system as the content feature of the items. The review text contains rich information of items. We denote each item x_i by a d-dimensional word vector. We denote the user-item rating matrix by $R € R^{nxm}$. The value in the matrix Ris rated by the users in the recommender systems, which indicates the user preference for the recommended items. We notice that the user ratings are sequentially collected and we denote that the collection of user ratings at the kth timestamp (or the kth round) by Ω_k .We use the term timestamp and round interchangeable throughout the paper. We then denote the collection of user ratings from the 1th round to the Kth round by $\Omega=\{\Omega_{1,...}\Omega_k,......\Omega_k\}$ The rating value of the jth item by the 1th user exists if its index exists at certain Round k.We notice that the matrix R is sparse and a number of values are missing in R. We now show that the learning of user preference by incorporating both user-item collaborative relationship and content feature of the items. We represent the rating prediction of the users by function $f_w(.)$ where vector w is the user preference.

4. Determination of Ratings Using Frank-Wolfe Algorithm

Social recommender system termed Online Graph Regularized User Preference Learning (OGRPL). In the task of online recommendation, the number of user ratings collected at each time stamp is much smaller than the ratings in the offline recommendation, which means all the items have to be recommended in a cold-start manner. Currently, social networking and knowledge sharing sites like Twitter and Douban are popular platforms for users to generate shared opinions for the items like item review and summary. Thus, the user generated content provides the auxiliary information for the items, which has been widely used to tackle the problem of cold-start item. Unlike the existing online collaborative filtering methods OGRPL is a hybrid model utilizing both CF information via the partially observed user-item matrix as well as the auxiliary content features for each item. Given a stream of user ratings, OGRPL incrementally learns the user preference on the content features of the items.

4.1 Collaborative Filtering

Collaborative filtering filters information by using the recommendations of other people, that bases its predictions and recommendations on the ratings or behavior of other users in the system. A person who wants to see a movie for example, might ask for recommendations from friends. The recommendations of some friends who have homogeneous interests are trusted more than recommendations from others.. A large number of collaborative filtering systems have to be able to handle many users .Making a prediction based on the ratings of people has implications for run-time performance. Therefore, when the number of users rating reaches a certain amount a selection of the best neighbors has to be made. Two techniques, correlation-thresholding and best-n-neighbor, can be used to determine which neighbors to select. The first technique selects only those neighbors whose correlation is greater than a given threshold. The second technique selects the best *n* neighbors with the highest correlation.In our project we concentrate on Model-based a collaborative filtering algorithm that provides item recommendation by developing a model of user ratings. In this category Algorithms take a probabilistic approach and envision the collaborative filtering process as computing the expected value of a user prediction, given his/her ratings on other items.

4.2 Online Optimization

In this section, we present a simple but efficient online optimization method to solve Problem. Given a sequential collection of user ratings at $\Omega=\{\Omega_{1,....,}\Omega_k\}$, the online optimization method aims to learn the user preference from $W_1,...W_k$, where the trace norm ball of each is bounded by γ (i.e.,$||W_k||_* \leq \gamma$) . That is, at each round k , the optimization learns the user preference W_k based on the current user ratings at and the previous preference matrix W_{k-1} .Suppose the optimal loss value of bath preference learning is $\min_{||Wk||_* \leq \gamma} F_k(W)$ and the user preference learned atthe Kth round is W_k,the goal of our online optimization method is to learn W_K is such that the regret,

$$\text{Regret}= F_k(w_k) -,$$

is sublinear in K,We first have a brief discussion on the property of the trace norm constrained objective function in Problem . We observe that this objective function falls into the general category of convex constrained optimization, which can be solved by Frank Wolfe's conditional gradient descent method. Convex constrained optimization refers to the convex optimization problem with a convex objective function and a compact convex domain.

4.3 Regularization Term

Regularization term is added to avoid overfitting problem. Overfitting have poor prediction and generalization power. It sticks too much to the data and the model has probably learned the background noise while being fit. To avoid this problem where the regularization technique comes in handy,

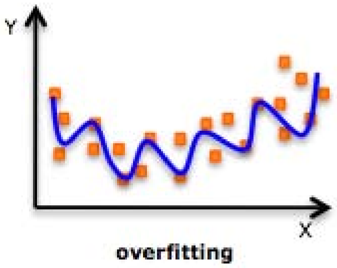


Figure 1: Overfitting

A simple linear regression is an equation to estimate y, given a bunch of x. The equation looks something as follows:
y = a1x1 + a2x2 + a3x3 + a4x4

In the above equation, a1, a2, a3 … are the coefficients and x1, x2, x3.. are the independent variables. Given a data containing x and y, we estimate a1, a2, a3 …based on an objective function. For a linear regression the objective function is as follows:

$$\min_{f} |Y_i - f(X_i)|^2$$

Now, this optimization might simply overfit the equation if x1, x2, x3 (independent variables) are too many in numbers. Hence we introduce a new penalty term in our objective function to find the estimates of co-efficient. Following is the modification we make to the equation :

$$\min_{f \in H} \sum_{i=1}^{n} |Y_i - f(X_i)|^2 + \lambda \|f\|_H^2$$

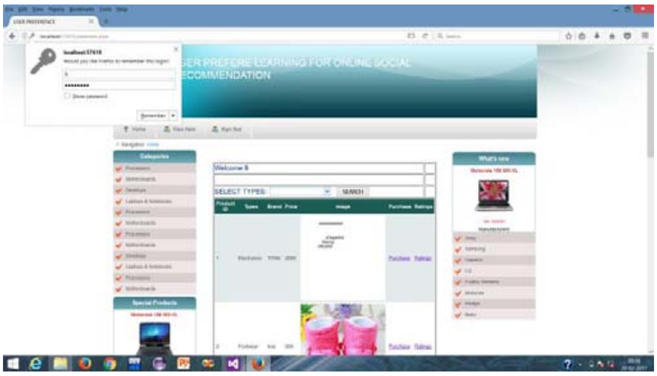
The new term in the equation is the sum of squares of the coefficients (except the bias term) multiplied by the parameter lambda. Lambda = 0 is a super over-fit scenario and Lambda = Infinity brings down the problem to just single mean estimation. Optimizing Lambda is the task we need to solve looking at the trade-off between the prediction accuracy of training sample and prediction accuracy of the hold out sample.

4.4 Buyer to Buyer Interaction

In this project chat facility is included for improving the interaction between the buyers. First of all, group the customers belonging to the same category.when the user logged into the site the recommended products and the product with highest rating is displayed first, the customer has to select the product to buy. Then the chat application group pops up, where the new user will be able to read the previous messages in the group to get the review.This application requires for the users to create a separate log in form, where a unique id is given to all the users. Either a group or a private chat can be created based on the user convenience

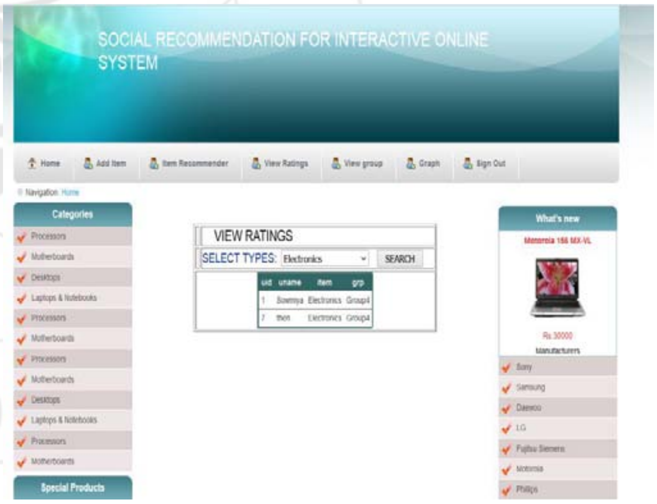
5. Experimental Results

5.1 Searching Items



5.2 View Rating

5.3 Grouping Customers



6. Conclusion

Online social recommendation which incorporates both collaborative user-item as well as item content features into a unified preference learning process. We consider that the user model is the preference function which can be learned from the user-item rating matrix. Furthermore, our approach integrates both user preference learning and user's social relations into a common framework for the problem of online social recommendation. In this way, our method can further improve the quality of online rating prediction for the missing values in the user-item rating matrix and chatting application for the customers. We devise an efficient iterative procedure, OGRPL-FW with Regularization Penalty term to solve the online optimization problem. We conduct extensive experiments on several large-scale datasets, in which the encouraging results demonstrate that our proposed algorithm achieves better performance than the state-of-the-art online recommendation methods.

Reference

[1] X. Yang, H. Steck, and Y. Liu, "Circle-based recommendation in online social networks," in Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2012, pp. 1267–1275.
[2] J. Zhu, H. Ma, C. Chen, and J. Bu, "Social recommendation using Low-rank semi definite program," in Proc. 25th AAAI Conf. Artif.Intell.,2011, pp. 158–163.
[3] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized recommendation combining user interest and social circle,"IEEE Trans.Knowl. Data Eng., vol. 26, no. 7, pp. 1763–1777, Jul. 2014.
[4] J. Tang, J. Tang, and H. Liu, "Recommendation in social media:Recent advances and new frontiers," in Proc. 20th ACM SIGKDD Int. Conf. Knowl.Discovery Data Mining, 2014, pp. 1977–1977.
[5] M. Ester, "Recommendation in social networks," in Proc. RecSys ,2013,pp. 491–492.
[6] H. Ma, "An experimental study on implicit social recommendation," in Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval , 2013, pp. 73–82.
[7] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King, "Recommender systems with social regularization," in Proc. 4th ACM Int. Conf. Web Search Data Mining , 2011, pp. 287–296.
[8] H. Gao, J. Tang, X. Hu, and H. Liu, "Content-aware point of interest recommendation on location-based social networks," in Proc. AAAI Conf.Artif. Intell., 2015, pp. 1721–1727.

Author Profile

Ms. V. Thenmalar is currently pursuing Bachelor's degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest includes Data mining and OOPS concepts.

Ms. R. Tamilselvi is currently pursuing Bachelor's degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest includes Data mining and Data Structures.

Ms. S. Sandhiya is currently pursuing Bachelor's degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest includes Data mining and OOPS Concepts.

Ms. M. Dhivya Shree is currently working as Assistant Professor in Department of Computer Science and Engineering. Her area of interest is Data Mining