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 4, April 2017

A Survey on User Preference Learning In Recommendation System

Jyoti Pardeshi, Prof.M.S.Takalikar

M.E. Student, Dept. of Comp Engineering, Pune Institute of Computer Technology, SPPU, Pune, India

Assistant Professor, Dept. of Comp Engineering, Pune Institute of Computer Technology, SPPU, Pune, India

ABSTRACT: The proposed survey discusses the user preference learning in the context of online e-Commerce websites. Preference learning helps to identify the preferences of a set of objects, by extrapolating known preferences of a similar, or possibly the same, set of objects. Various e-Commerce websites maintain user related information according to user provided data and tracking it's activity on specific website. Due to complexity in building user profile; preference learning is challenging task. In this survey various preference learning are discussed along with applicability of these approach.

KEYWORDS: Recommendation system, user preference learning, online social networks

I. INTRODUCTION

E-commerce websites from various sectors e.g. marketing (Amazon/Flipkart), entertainment (Netflix), Search & Discovery etc. become an integral part of people day-to-day activity due to ease in finding useful information about interested things quickly in cheaper manner. Ease in using e-commerce application on mobile devices achieves rapid growth in user base and leads to generate vast amount of information them. Statistical study of popularity of e-commerce websites results in following facts:

- 80% of population which is on internet purchased something from e-commerce websites(Source: Custora's E-Commerce Pulse)
- Before purchasing any commodity product user spend on 16-30 minutes an average on comparison of products.

As e-commerce applications provides relief to user from getting overloaded by information about products; consumer base of such applications is increasing day by day and generates huge data through user interaction/ browsing and transaction activity.

Gathering user characteristics to build profile, identify interest and/or preference and finally provide recommendation accordingly is important and challenging task to achieve user satisfaction. Though amount of users purchasing an items online is large, purchase/user is relatively small. Along with that finding interest of naïve users/ users who does only comparison shopping is tedious task. Identifying such different scenarios and provide recommendation accordingly taken care by preference learning.

II. LITERATURE SURVEY

In this section survey of preference learning approaches is presented. Further section provides classification of those techniques along with datasets on which techniques are applied. We have excluded techniques based on utility functions as it is not part of our scope.

1. Model based learning

Luiz Pazzato et al[3] suggested preference learning approach for those social networking websites where user's generally hesitate to provide explicit feedback e.g online dating sites. Important distinction in recommendation systems for such approach is both parties e.g. User/Item are active participants in achieving recommendation. In proposed

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 4, April 2017

technique interest identification is carried out by tracking user actions on website. Ranking approach considers reciprocal compatibility score.

Akehurst J. et al[4] performs preference learning for online dating website only but interest identification is carried out on the basis of interaction using static text messages. Score (Positive/Negative) is associated with each text message. Finally precision score of the messages in total interaction is considered for recommendation. Though it is simpler approach interest identification technique is too rigid and could not scale well.

Though above techniques used for two-side preference learning those are limited to single application e.g. online dating. Anjan Goswami et. al[6] proposes generalized regression model based preference learning which scales to variety of two-side preference learning markets. Proposed model works in two phases. In first phase for both side participants probabilities are identified using regression model. In second phase cross validation technique is used to tune regularization parameter. Experimental results shows improvement is AUC.

2. Memory based learning

Dhiraj Goel et al[5] evaluate preference learning approach for movie recommendation website. In proposed approach cold-start problem tackled with the help user clustering using k-nearest neighbor algorithm. Finally average response of each cluster is used to predict rating of movie which is not seen by user.

3. Preference relations:

S Liu et al[7] proposed technique for recommendation using preference relations. In proposed approach rating is calculated with the help of relative preference of items made by user. Later markov random field is used to identify item-to-item user preference relation. Finally ranking achieved using regression approach.

Alan Echardt[8] studied the preference learning problem in context of fuzzy logic. According to author two types of preference e.g. local and global are evaluated with the help of features of items. Finally based on global score evaluated per item will be used to identify ranking.

To leverage social relationship associated with users Felicio[9] proposed pairwise preference recommendation model. Proposed technique evaluate user connection weight on the basis of factors like friendship, interaction level and mutual friends etc to tackle new user recommendation problem. Learning framework works in two phases. Process initiates with clustering user's according to user-item rating matrix and later average rating of each cluster is calculated known as consensus calculus.

Sr No	Approach	Dataset	Evaluation measure	Pros &Cons
1	Histogram based model learning [3]	Not Disclosed	Precision &Recall	*Pros*: Beneficial in case of lack in user feedback *Cons*: Active participation of both parties expected for recommendation
2	Histogram based model learning [4]	Australian dating website <Name not disclosed>	Precision &Recall	*Pros*: Same as above *Cons*: Approach used for interest identification is too rigid.
3	Memory based learning [5]	Netflix	RMSE (Root Mean Square Error)	*Pros*: User clustering prefers item preference score over static friendship link score *Cons*: Computational cost is high with growth of data
4	Linear regression model learning [6]	Elance.com	AUC (Area Under the receiver	*Pros*: Proposed applicable model for variety of two sided social networking sites

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 4, April 2017

			operator Curve)	
5	Preference relation based learning [7]	MovieLens, Eachmovie	MAP (Mean Average Precision	*Pros*: Preference relation information studied in the context of Markov Model Field which is beneficial in case of real time prediction *Cons*: Prediction are made only basis neighbours
6	Preference relation based learning [8]	Movielens	RMSE, Tau coefficient	*Pros*: Normalization of item attribute values while calculating local preference compare attributes on equal measure and reduces complexity *Cons*: No experimental evaluation provided which will helpful support effectiveness of approach
7	Preference relation based learning [9]	Facebook, Flixter	Precision &Recall	*Pros*: Preference matrix replaced with calculus which reduces complexity and execution time *Cons*: New user recommendation problem persists if user has less number of friends

IV. Related work

Recommendation system is nothing but information filtering system which provides solution to Information Overload problem by analyzing user interest and caching those details to provide relevant suggestions in future [1]. Following figure depicts the process in typical personalized recommendation systems.

1. Generated user profile: Store details about the user
   E.g: Basic Information
2. Maintain user profile: Update profile details according to user actions and feedback.
   E.g. Ratings provided, Items searched & crawled
3. Exploit user profile: Use profile data for recommendation.
   E.g. Purchase history, neighbor relationship

Though above process followed by most of the algorithms, use of recommendation approach is depends upon various factors. Following section describes the details about recommendation approach and applicable scenarios of it.

A. Recommendation approaches:

With emergence of need of recommendation systems in various sectors various approaches are introduces. Following factors plays important role in defining any recommendation approach:

a. Dataset :

Data variety e.g. information about various objects and relationship between them and size .e.g. amount of information available are important in deciding accuracy and performance of recommendation.

b. Data description :

Data content either be qualitative e.g. objects with attribute details or quantitative e.g. integer scale based or both. Nature of data defines characteristics and limitations in user profile modelling. According to above factors following classification of well-known recommendation approaches is provided.

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

International Journal of Innovative Research in Computer and Communication Engineering

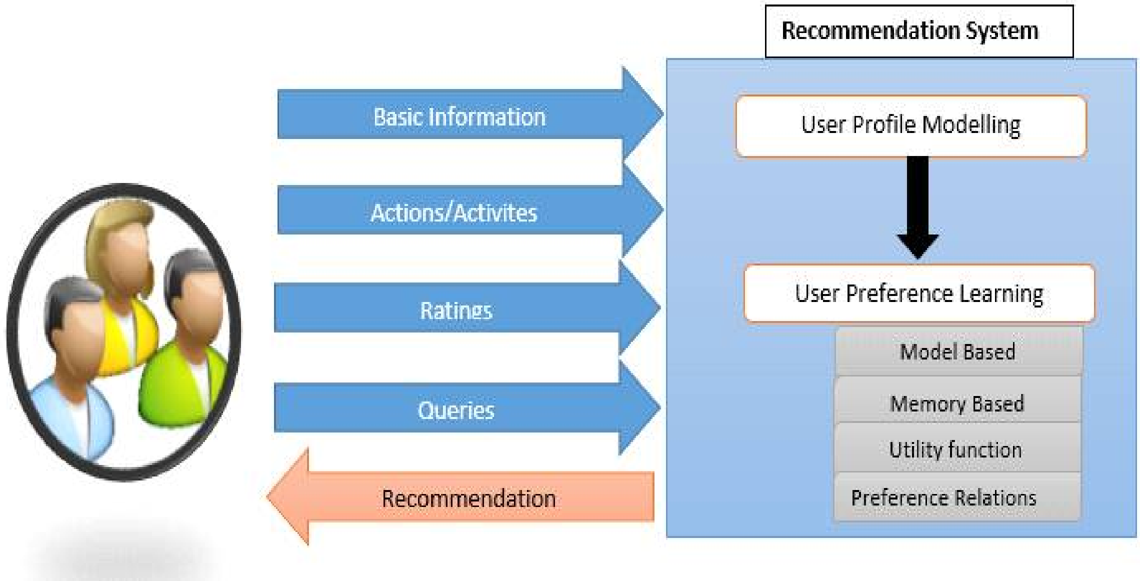
(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

Sr. No	Recommendation Approach	Description	Properties *(Dataset & Data description)*
1	Content based filtering	Provides the recommendation on the basis of attributes of items rated by user in past	Qualitative item information expected
2	Collaborative filtering	Provides the recommendation on the basis of users who share similar interest	Quantitative user-item rating information expected
3	Hybrid	Combination of above	-
4	Knowledge-based	Predicts/Infers content according to knowledge of domain and user actions	Qualitative user action information expected

In each of the above approach modelling user profile (Maintain User Profile) and learning user preference (Exploit User Profile) are important phases. User modelling is process collecting facts which defines any individual on recommendation system while learning user preference is to identify interest of individual based on available relevant history.



B. User Profile Modelling:

In process of recommendation user does some action with recommender system[1][2]. Based on those activities user profile is generated which will later analyzed in preference learning phase. To model user profile some following information needs to be processed.

- Implicit Feedback

Implicit feedback is evaluated by tracking user activities and interpretation of it. Sources of implicit feedbacks are different types of behaviors e.g. Click-throughs, crawling to/skipping particular web page, waiting time. Main advantage of implicit feedback it avoid relevance judgement by user and it can be used to provide confidence while evaluating user interest. In case of lack implicit feedback explicit feedback are also used to calculate pseudo implicit feedback.

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 4, April 2017

•Explicit Feedback

Explicit feedback is gathered through acquiring knowledge about specific object from user. Feedbacks can be gathered as qualitative or quantitative approach. E.g. Questionnaire provided to user for gathering rating scaled in range or categorization into different classes e.g. Like/Dislike. Explicit feedbacks are important in case of determining preference of user interest. Though explicit feedback are easy to collect and free from complex processing for interpretation, reliability of such feedback is one of the important issue.

Based on nature of data collected e.g. structured/unstructured involvement of post-processing task identified. Later those data maintain in appropriate data structure e.g. Matrix/Tree so that it will be easily used for preference learning.

C. User Preference Learning:

Learning task comprises analysis of information associated with user model. Preference learning approaches are categorized based on following parameter:

A. Computation strategy:

Algorithm evaluation carried out in online or offline environment based on characteristics of user profile data. Online evaluation also known as Memory based/ Heuristic based learning which is used in real time recommendation as it used in case of frequent change observed in user interest. While offline evaluation aka Model based approach builds the model which is time consuming approach and preferred if change is user behavior is slow.

B. Recommendation results evaluation

Recommendation results are provided in two different ways. “Partial ordering” where recommendation list is presented to users based on overall interest of users irrespective of whether ordered objects are belongs to same domain or not. While “Total ordering” take care of ordering objects according to domain which is more qualitative approach. Applicability of this approach is totally depend upon nature of data.

There are four types of preference learning approach described below:

1. Model based learning:

It is common approach to learn preference [1][2][6]. Process of learning preference starts with building classifier from training set. Once the classifier is built it is able to classify/label new incoming input system. Classifier is implemented using machine learning techniques e.g. probabilistic approach, decision trees, neural networks, Bayesian networks and association rules.

2. Memory based learning:

Memory-based approach work on the principal of aggregating the labeled data and attempt to match recommenders to those seeking recommendations [1][5]. Most common methods are based on the notion nearest neighbor metrics. Methods of this approach utilizes whole dataset and scale linearly with size of data.

3. Utility functions

Utility based methods rank objects using a utility function which represents the degree of usefulness of the object [2]. There are different methods in implementing utility methods for label ranking, object ranking and instance ranking. Generally utility score is assigned to each object where utility scale is either numerical or ordinal. Based on assigned score goal is to improve ranking performance

4. Preference relations:

Key idea of this approach is based on provided comparative information about preference of objects find order of interesting objects [2][6]. Pairwise preference of objects are collected and IR techniques are used to identify proper ranking of objects. Main advantage of this approach is few or no interpretation of training data is required. Most of the content based recommender systems used this approach as variety of nominal/non-ordinal features/attributes about objects are available for comparison.

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 4, April 2017

Most important factor which is considered in approach selection is availability of input sources. In real time application memory based and preference relation approaches are most preferred due to adopting dynamic changes and provision of most useful recommendation.

V. PROPOSED APPROACH

Methodology is proposed for improving recommendation process by reducing time required to process dynamic and vast amount of available data [10]. It will remain useful for OSNs which are collecting user reviews/ratings and providing recommendation on users historical data e.g. Douban, CIAO. [11] In proposed approach, our main goal is to provide user preference matrix in time bounded manner and improve performance of preference learning process.Main phases of User preference learning system are explained in detail as follows:

A. Matrix Generation

Using datasets User rating matrix R, Item content feature (Item-Category) matrix X , User similarity matrix S and User Preference matrix W are formed. In this phase R, W matrices are very sparse due to missing values of user rating for various items. User similarity matrix calculated using Jaccard Distance .

B. Rating matrix based loss value detection

User-Item rating matrix R can provide details about exact preference of users for specific time interval but over the time user preferences varies. At k-th time interval, user preference W k calculated with the help of current rating matrix Ω k and previously learned {W 1 ,...,W k−1 } using {Ω 1 ,...,Ω k−1 }. Rating values are provided in numeric range where user choice is one of the probable value in range[10]. To avoid the hard constraint of noisy rating values provided by user the loss function is used[12].

C. Relationship matrix based loss value detection

According social homophily, user relationship network also leads to interest propagation[13]. This factor helps in reducing data sparsity in user-rating matrix by detecting community of user's in social network. Various social network is directed graph e.g. Twitter where link may be unidirectional and/or bidirectional. Unidirectional link depicts weak relationship in such network which restricts interest propagation. To gain advantage of social relationship by removing dependency of weaker relationship the loss detection formula is used.

D. User preference calculation using frank wolf algorithm

To predict future user preference matrix, we have formulated loss function formula in previous stage. fter applying loss detection formula resultant user preference matrix will contain set of values whose bounding range depends upon user-item rating history [14]. Such set of values known as convex set. To solve convex set problem, frank-wolf algorithm; a conditional gradient descent method. sing frank wolf algorithm user preference matrix loss function formula is used[14].

VI. CONCLUSION

This is an overview of the preference learning process in recommendation systems. The existing approaches are illustrated with the main focus on input sources used in preference learning. Type of dataset affects the user modelling approach and results in improving personalized recommendation.

REFERENCES

1. M. De Gemmis, L. Iaquinta, P. Lops, C. Musto, F. Narducci, and G. Semeraro. Preference learning in recommender systems. In ECML/PKDD-09 Workshop on Preference Learning, 2009
2. J. Fürnkranz and E. HÄullermeier. Preference learning, Kunstliche Intelligenz, vol. 19, no. 1, pp. 60-61, 2005
3. L. , In Proc. Preference Learning Workshop, ECML PKDD Conference, 2010

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 4, April 2017

4. Akehurst, J., Koprinska, I., Yacef, K., Pizzato, L.A.S, Kay, J. and Rej, T. Explicit and Implicit User Preferences in Online Dating. International Workshop on Behavior Informatics, Lecture Notes in Computer Science 7104, Springer, pp. 15-27, 2012
5. D. Goel and D. Batra. "Predicting User Preference for Movies using NetFlix database." Department of Electrical and Computer Engineering, Carniege Mellon University, 2009
6. A. Goswami, F. Hedayati and P. Mohapatra, Recommendation Systems for Markets with Two Sided Preferences, 13th International Conference on Machine Learning and Applications, Detroit, MI, pp. 282-287, 2014
7. S. Liu, G. Li, T. Tran and Yuan Jiang, Preference Relation-based Markov Random Fields for Recommender Systems, In Proceedings of Machine Learning Research, Vol. 45, pp. 157-172, 2016
8. A. Eckhardt. Various aspects of user preference learning and recommender systems. In DATESO, pages 56–67, 2009.
9. A. Felício, K. Paixão, G. Alves, S. De Amo, P. Preux. Exploiting Social Information in Pairwise Preference Recommender System. Journal of Information and Data Management, Brazilian Computer Society, Vol 7 (2), pp.16, 2016
10. Z. Zhao, H. Lu, D. Cai, X. He and Y. Zhuang, "User Preference Learning for Online Social Recommendation," in IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 9, pp. 2522-2534, Sept. 1 2016
11. M. Jaggi. Revisiting frank-wolfe: Projection-free sparse convex optimization. In Proceedings of the 30th International Conference on Machine Learning (ICML-13), pages 427435, 2013.
12. Z. Zhao, L. Zhang, X. He, and W. Ng. Expert finding for question answering via graph regularized matrix completion. Knowledge and Data Engineering, IEEE Transactions on, pages 9931004, 2015
13. X. Ding, X. Jin, Y. Li, and L. Li. Celebrity recommendation with collaborative social topic regression. In Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, pages 26122618. AAAI Press, 2013.
14. C. Luo, W. Pang, and Z. Wang, Hete-cf: Social-based collaborative filtering recommendation using heterogeneous relations, arXiv preprint arXiv:1412.7610, 2014

BIOGRAPHY

Jyoti Pardeshi is a student of M.E. in the Computer Engineering Department, Pune Institute of Computer Technology. Her research interests are Data Mining, Information Retrieval.

M. S.Takalikar is a assistant professor Department of Computer Engineering, Pune Institute of Computer Technology. Her research interests are NLP, Machine Learning, Data Mining.