International Journal of Pure and Applied Mathematics
Volume 117 No. 22 2017, 297-301
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

PEROSONALIZED RECOMMENDER SYSTEM FOR MINING THE REAL TIME WEB

Janet Rajeswari[1] , Shanmugasundaram Hariharan[2]
[1]Research and Development Centre, [2]Department of Information Technology
[1]Bharathiar University, Coimbatore- 641046, India, [2]Vel Tech Multitech Engineering College, Chennai, India
e-mail: janetcharles1@hotmail.com,mailtos.hariharan@gmail.com

Abstract: The tendency of analyzing user reviews has been increasing tremendously due to the inherent nature of online web users. The scenario of buying/selling products online with a hassle free environment has attracted people of all age groups. The commercial web sites attract people through various ways advertisements and offers and increase their online marketing. There lies some negative impact in introducing fake products, poor recommendation by offenders and others. Therefore recommender systems are of greater importance in providing valuable reviews to end users for online buyers/sellers. We present a hybrid product recommender system with the integration of multiple features to mine the real time Web 2.0 applications. Experiments performed show the significance of the proposed approach in terms of standard evaluation metrics.

Keywords: Recommender systems, precision, recall, personalization, information retrieval.

1. INTRODUCTION

Several Web 2.0 applications allow users to assign keywords (or tags) to provide better organization and description of the shared content. With the explosion of Web 2.0 applications such as twitter, blogs and social media, information sharing has become an easy task. However these online information which have emerged from various online e-sources pose a great challenge in terms of information validation. There are millions of users who spend several hours of time on day to day basis on these commercial sites and generate information from various sources of knowledge with item specific tag preferences [1]. Thus the information available is used for right sense by the online vendors and people to make use of the information to get recommendation from different people. The abundance and popularity of social networking sites flood users with huge volumes of information and hence pose a great challenge in terms of information overload and also create many new research avenues for variety of recommender systems [11]. Recommender systems have originated over a decade with investigations on recommendation systems for book recommender process using genetic approach [12].

Recommender systems have variations with respect to social web storage, search and mining process, social network construction and association with social networks etc. It also introduces many real world applications which operate in a full fledged manner with the help of these agents [10]. For examples, web community detection and search, hot-topic detection in a specific web community, accurate and timely recommendations in commercial applications. These research issues have been receiving growing attentions in recommender systems field, data mining and among others in the recent years. This paper presents a framework on to provide a recommendation algorithm investigating some of the attributes relating to recommendation process in a commercial context [8, 9]. Many significant researches have been done on these research topics. The main purpose of this document is to provide a survey of the development on these research challenges. Some of the recommender projects in existence includes CAWICOMS [2], PRemiSE [3], WebPUM [4] and PerHSS [5]. Recommender systems includes agents which are user-centric [7], a variety of different roles and responsibilities of recommender systems is also discussed [13]. It is indeed these social networks operate on a strong user centric model with more personalized behavior [6]. Recommender systems are broadly categorized as content-based methods, collaborative methods and hybrid methods. Some of the popular recommender systems are NewsWeeder [14], YourNews [15] and NewsDude [16].

The rest of the report is organized as follows. Section 2 briefs about some of the literature review carried out on recommender systems, work relating to different recommender systems. In Sect. 3, we present our recommender systems and the experimental results carried out using commercial datasets. Section 4 concludes the study with directions for future work on recommender systems.

2. LITERATURE REVIEW

This section briefs on some of the work investigated by several researchers in the area of recommender systems. Recommendation process has prone to be successful due to inherent nature of real time information availability, characteristics and several other factors associated with it. Of these several types of recommender systems, news recommenders have predominant in providing services to users obtained from popular news websites. The authors have reported on utilizing the social factors (i.e., the potential influential experts in news reading community) among news readers to facilitate news personalization by integrating content-based methods, collaborative filtering and information diffusion models employing probabilistic matrix factorization techniques temred as PRemiSE [17]. Such Personalized news Recommendation framework with implicit feedbacks was designed, evaluated and compared with various baselines on a collection of news articles.

Recommender tasks adopting Machine learning (ML) and data analytics tasks in practice require several consecutive processing steps. In the current era, RapidMiner is widely used as a software tool for the development and execution of such analytics workflows. Unlike many other toolkits available for public use, there comprises a visual editor allowing the user to design processes on a conceptual level [18]. These conceptual and visual approaches help the user in abstracting from the technical details during the development phase. RapidMiner framework supports the user in modeling phase by recommending additional operations to insert into the currently developed machine learning workflow. First, a recommendation technique is used to evaluate in an offline setting using several thousand existing workflows. Subsequently, the tools help the users in significantly increasing the efficiency of the modeling process. RapidMiner is a software tool that allows users to define data mining processes based on visual model and implements variety of so-called “operators” for data extraction, manipulation, model learning and analysis [21]. The large number

of available operators can however make it challenging for the process designer to find the appropriate operators for the problem at hand. At the same time, some operators are only meaningful when combined with certain others. In this work, we evaluate different strategies of recommending additional operators to the user during the design of the process. The recommendation models are learned using a pool of several thousand existing data mining processes and evaluated in an offline experiment. The results indicate that good predictive accuracy can already be achieved with comparably simple co-occurrence based algorithms.

The recommender systems rely strongly on trust model and are considered to be a key success factor for automated recommendation agents on e-commerce sites. Various aspects contribute to the development of trust toward such an agent, including perceptions about the usefulness of the recommendations, the transparency of the recommendation process, and the general quality of the website [19]. A new model was proposed and evaluated that integrates these factors which allows to assess the relative importance for trust-building. The empirical studies in the context of two popular e-commerce websites, shows that the findings observed were transparent and is equally important to consumers for building trust as recommendation quality. It is also inferred that the recommendation quality may be good enough at lower levels and insufficient as far as higher levels of adoption of the recommendations is concerned. This result in analyzing and building up several trust factors.

Online shops and B2C (business to consumer) sites in diverse domains such as 'quality & taste', consumer electronics or e-tourism require persuasive web presentation on one hand and deep product knowledge at other end. These context recommender applications help in creating an enjoyable shopping experience for online users. The designed Advisor Suite framework deals with knowledge based conversational recommender system that aims at mediating between requirements and desires of online shoppers and technical characteristics of the product domain. The work presents a conceptual scheme to classify the driving factors for creating a persuasive online shopping experience with recommender systems [20]. The work also discusses on the concepts based on several fielded applications.

The conceptual framework using multimodal user feedback generates more accurate personalized ranking of items to the user. The proposed technique is indeed a response to the actual scenario on the Web, where the users can consume content following different interaction paradigms, such as rating, browsing, sharing, etc. The work has developed a post- processing step to ensemble rankings generated by unimodal based state-of-art algorithms, using a set of heuristics which analyze the behavior of the user during consumption. The experimental evaluation using the MovieLens 10M dataset[# $], illustrates that the results show better recommendations which could be provided when multimodal interactions are considered for profiling the preferences of the users [22].

Web 2.0 enables tag recommendation methods which assist users in improving the quality of the available information and thereafter the effectiveness of various tag-based information retrieval services, such as searching, content recommendation and classification. Some authors have addressed tag recommendation problem in two different perspectives namely suggesting relevant tags to a target object by exploiting three dimensions as tag co-occurrences, terms extracted from multiple textual features (e.g., title, description) and estimating tag relevance. The second perspective aims at performing personalized tag recommendation to a target object user pair, exploiting, three dimensions discussed in first perspective [23]. A new heuristic method with the state-of-the-art strategies including new metrics is estimated to determine the accuracy of target object generation. The authors exploited three learning-to-rank (L2R) based techniques, namely, RankSVM, Genetic Programming (GP) and Random Forest (RF), for generating ranking functions that exploit multiple metrics as attributes to estimate the relevance of a tag to a given object or object-user pair. The proposed methods were evaluated using four popular Web 2.0 applications, namely, Bibsonomy, LastFM, YouTube and YahooVideo. The results were noted against recall, Normalized Discounted Cumulative Gain (NDCG) and Mean-Reciprocal Rank (MRR) which shows the significance of the proposed investigation [23].

The emergence of Internet and popularity of online trading has attracted more and more customers in exchanging products on online stores. This has also raised a situation certain difficulty in selecting trustworthy and suitable seller who e-business. Work done in recent years [24] proposed a personalized trustworthy seller recommendation system for the customers of an open market in Korea. The authors have developed a module which classifies sellers into trustworthy one or not using a classification technique such as decision tree. Later a module which makes use of the content-based filtering method to find best-matching top k sellers among the selected trustworthy sellers is developed. Experimental results have shown that approach is significant and worthwhile.

Increase in the growth of e-commerce results in users having different needs, situations and contexts. It is increasingly important to consider user context data when filtering these vital information's. In the field of web personalization and recommender systems, most of the studies have focused on the process of modeling user profiles and the personalization process in order to provide personalized services to the user, but not on contextualized services and limited attention has been paid to investigate how to discover, model, exploit and integrate context information in personalization systems in a generic way [25]. To alleviate such issue, the authors aimed at providing a novel model to build, exploit and integrate context information with a web personalization system.

[#]http://www.grouplens.org/,[$]http://recsyswiki.com/wiki/Datasets

A Context Aware Personalization System (CAPS) is therefore developed which is able to model and build contextual and personalized ontological user profiles based on the user's interests and context information. These profiles are then exploited in order to infer and provide contextual recommendations to users. The methods and system developed by the authors are evaluated through a user study which shows that considering context information in web personalization systems can provide more effective personalization services and offer better recommendations to users.

In recent years information analysis and knowledge gathering applied through semantic network is applied to more and more crucial research areas, such as recommender system [26]. Differing from traditional users' recommendations, the tweets' recommendation in a micro-blog network has two crucial differences. One is high authority users or one's special friends usually play a very active role in tweet-oriented recommendation. Micro-blog user enables the users his/her very interested into ''special attention'' group, and the topics discussed more in ''special attention'' group are more likely to be the user interested topic. The other is that users hope to obtain more relevant tweets about what he/she is interested in. Thus, this work presented [26] uses the k-cores analysis method to extract topics that users pay attention to, and employs the method of factor analysis to analyze index, and to extract the tweet heat factor and user authority factor. Besides, the investigations also intend to use the method of RS and linear regression to determine the parameters for balancing the value of the tweet heat factor and user authority factor. Finally, this work manages to establish a timely personalized recommendation model

based on semantic network for SINA tweets. The experimental results, reveals the proposed method could effectively solve problems existing in micro-blog tweets in a personalized and timely recommendation way.

3. PROPOSED RECOMMEDNER SYSTEM

The system architecture (Fig .1), experimental investigations, results and descriptions is presented in this section. The user extracted opinionated reviews are analyzed in correlation with the user preferences. The learning agent scores the reviews with additional weights to terms related to the specific portion of user preferences. The trained set of reviews with user preferences is analyzed by the agent. The review ratings obtained MovieLens 10M dataset, consisting of 10 million ratings, 100,000 interactions tags applied to 10,000 users and 72,000 movies.

Table 1 presents the confusion matrix for product recommendation process. The user agreement among the users is measured using equation 1.These measures are calculated by considering the number of items which are either relevant or irrelevant and either recommended. Equation 2 & 3 presents the evaluation metric Precision and recall, which is widely used in the field of information retrieval, to evaluate recommendation accuracy. These metrics have been adapted to evaluate the accuracy of a set of recommended products and are defined as:

$$\cos ine(d_i, d_j) = \frac{|d_i \cap d_j|}{|d_i| * |d_j|} \quad (1)$$

$$\Pr ecision = \frac{|T \cap R|}{|R|} \quad (2)$$

$$\mathrm{Re}call = \frac{|T \cap R|}{|T|} \quad (3)$$

where 'T' is the test set and 'R' is the recommended set of items for each user, respectively. F1-metric can be used to balance the trade-off between precision and recall.

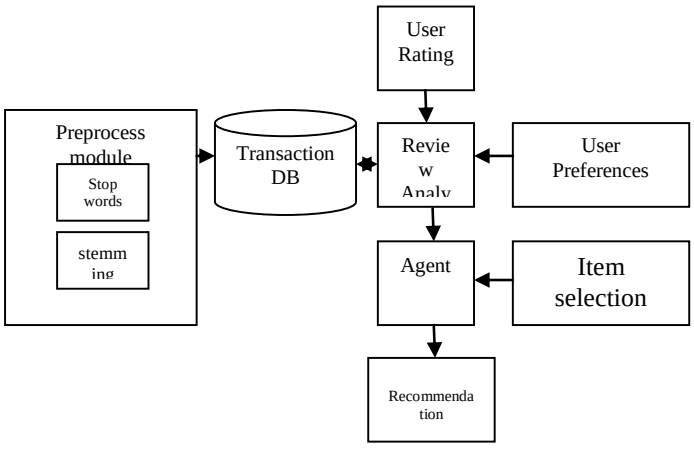

Figure 1. Architecture for recommender system

The framework identifies a set of recommended items (ranked by the active online users). The top ranked list of items generated by each user is extracted using association rules and regrouped depending the agreement among users. In FP- tree generation, the minimum support value is set as 20%. The performance of the proposed algorithm with other compared methods via different values of N. We performed a statistical test with a significance level of 95%. The *p*-values of the *t*-test statistics were computed and are shown in the following table 2. Singular Value Decomposition (SVD) have been used to compared with the proposed approach. Several other variants have been used with time related constraints which is presented for future studies.

TABLE 1: CONFUSION MATRIX

Evaluation metric	Approach	*Top percentile*				
		5	***10***	***15***	***20***	***25***
Precision	SVD	0.086	0.934	0.112	0.125	0.168
	Proposed	0.122	0.132	0.144	0.167	0.210
recall	SVD	0.028	0.042	0.081	0.105	0.121
	Proposed	0.087	0.075	0.123	0.187	0.205

TABLE II. PRECISION & RECALL OF PROPOSED APPRAOCH WITH SVD

Predictor	*Relevant*	*Irrelevant*
Recommended	T	F
Not recommended	F	F

4. CONCLUSION AND FUTURE WORK

This paper presented a study on recommender system task using MovieLens dataset using rule mining. Experimental illustration and the results presented is competitive in terms of precision and recall. The proposed work is compared with SVD which shows the significance of our proposed approach. Work focusing on generic time constrained recommendation task is under progress.

ACKNOWLEDGMENT

The authors would like to express our sincere thanks to anonymous reviewers for their insightful suggestions. The authors indeed would extend their acknowledgement for the support by all well wishers, faculty members and others who helped out in achieving this research task to get completed.

REFERENCES

[1] Gedikli, F., Jannach, D.: Improving recommendation accuracy based on item-specific tag preferences, ACM Transactions on Intelligent Systems and Technology 4(1), 2013.

[2] Ardissono L., Felfernig A., Friedrich G., Jannach D., Schäfer R., Zanker M.: Customer-Adaptive and Distributed Online Product Configuration in the CAWICOMS Project. In: M. Aldanondo (Ed.): 17th International Joint Conference on Artificial Intelligence - Configuration Workshop. Lyon: August 2001, pp. 8-14.

[3] Chen Lin, Runquan Xie, Xinjun Guan, Lei Li and Tao Li, "Personalized news recommendation via implicit social experts", Information Sciences,Vol. 254, pp. 1–18, 2014.

[4] Mehrdad Jalali, Norwati Mustapha, Md. Nasir Sulaiman and Ali Mamat, "WebPUM: A Web-based recommendation system to predict user future movements", Expert Systems with Applications 37 (2010) 6201–6212.26.

[5] Donghee Yoo, "Hybrid query processing for personalized information retrieval on the Semantic Web", Knowledge-Based Systems 27 (2012) 211–218.

[6] Felfernig, A., Friedrich, G., Isak, K., Shchekotykhin, K., Teppan, E., Jannach, D.: Automated Debugging of Recommender User Interface Descriptions, Applied Intelligence (Best papers from IEA/AIE 2006), Vol. 31(1), Springer, pp. 1-14

[7] Jannach, D., Lerche, L., Jugovac, M.: Item familiarity as a possible confounding factor in user-centric recommender systems evaluation, i-Com Journal For Interactive Media, 14(1) 2015, 29-40.

[8] Gedikli, F., Jannach, D., Ge, M.: How should I explain? A comparison of different explanation types for recommender systems.

International Journal of Human Computer Studies, Vol 72(4), 2014, Springer, pp. 367-382.

[9] Zanker, M., Jessenitschnig, M., Jannach, D., Gordea, S.: Comparing recommendation strategies in a commercial context, IEEE Intelligent Systems, Special issue on Recommender Systems, Vol. 22(3), 2007, pp. 69-73.

[10] Nilashi, M., Jannach, D., Dalvi, M., Ibrahim, O., Ahmadi, H.: Recommendation Quality, Transparency, and Web Site Quality for Trust-Building in Recommendation Agents, E-Commerce Research and Applications, Vol. 19, Sep/Oct 2016, pp. 70-84.

[11] Laura Cruz Quispe; José Eduardo Ochoa Luna, "A Content-Based Recommendation System Using TrueSkill, In proceedings of Fourteenth Mexican International Conference on Artificial Intelligence (MICAI), 2015, pp.2013-207, 2015.

[12]Celalettin Aygün and Oktay Yıldız, "Development of content based book recommendation system using genetic algorithm", In proceedings of 24th Signal Processing and Communication Application Conference (SIU), Pages: 1025 – 1028, 2016 .

[13] Aiswarya Thomas; A. K. Sujatha, "Comparative study of recommender systems", In proceedings of International Conference on Circuit, Power and Computing Technologies (ICCPCT), pages:1-6, 2016.

[14] NewsWeeder [17] – K. Lang, "Newsweeder: Learning to filter netnews," in Proc. of the 12th Int. Conf. on Machine Learning, 1995, pp. 331–339.

[15] YourNews [13]- J. Ahn, P. Brusilovsky, J. Grady and D. He., "Open user profiles for adaptive news systems: help or harm?" in Proc. of the 16th Int. Conf. on WWW, 2007, pp. 11–20.

[16] NewsDude [7]- D. Billsus and M. Pazzani., "A hybrid user model for news story classification," in Proc. of the 7th Int. Conf. on User Modeling, 1999.

[17] C.Lin, R.Xie, X.Guan, L.Li and T.Li., "Personalized news recommendation via implicit social experts", Information Sciences,Vol. 254, pp. 1–18, 2014.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[18] Dietmar Jannach, Michael Jugovac and Lukas Lerche, "Supporting the Design of Machine Learning Workflows with a Recommendation System", ACM Transactions on Interactive Intelligent Systems (TiiS), Volume 6 Issue 3 (November 2016)

[19] ehrbakhsh Nilashi, Dietmar Jannach, Othman bin Ibrahim, Mohammad Dalvi Esfahani and Hossein Ahmadi, "Recommendation quality, transparency, and website quality for trust-building in recommendation agents", Electronic Commerce Research and Applications 19 (2016) 70–84.

[20] Zanker M., Bricman, M., Gordea, S., Jannach, D., and Jessenitschnig, M., Persuasive product recommendation, in: Filtering, A., Zanker, M. (eds.): Workshop on Recommender Systems at ECAI 2006, Riva del Garda, Italy, 2006, pp. 6-10.

[21] Jannach, D., Fischer, S.: Recommendation-based modeling support for data mining processes, ACM Recommender Systems (RecSys 2014), Foster City, CA, pp. 337-340.

[22] Arthur F. Da Costa, Marcos A. Domingues, Solange O. Rezende and Marcelo G. Manzato, WI-IAT '14 Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 01 ,Pages 198-204 .

[23] U-Fabiano M. Belém, Eder F. Martins, Jussara M. Almeida and Marcos A. Gonçalves, "Personalized and object-centered tag recommendation methods for Web 2.0 applications", Information Processing and Management, Vol. 50, pp.524–553, 2014.

[24] Seungsup Lee, Keunho Choi and Yongmoo Suh, "A personalized trustworthy seller recommendation in an open market", Expert Systems with Applications 40 (2013) 1352–1357.

[25] Ahmad Hawalah and Maria Fasli, "Utilizing contextual ontological user profiles for personalized recommendations", Expert Systems with Applications, Vol. 41, pp. 4777–4797, 2014.

[26] Yue He and Jinxiu Tan, "Study on SINA micro-blog personalized recommendation based on semantic network", Expert Systems with Applications, Vol. 42, pp. 4797–4804, 2015.