Volume 6, Issue 6, June 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

K-Nearest Neighbor (k-NN) based Collaborative Learning for Multivariate Product Ranking on Shopping Portals

Manpreet Kaur*, Pravneet Kaur
CGC Technical Campus, Jhanjeri, Mohali,
Punjab, India

Abstract— The product ranking models play the vital role in the e-commerce shopping portals, as they are responsible for viewing the posts over the product page. The product ranking lists are created and converted into the product lists in order to facilitate the users for the easy navigation, as well as to increase the per user access reach over the e-commerce portals. The well maintained product list is responsible for the increase in the sales, when evaluated with the popular and compatible products to attract the more users and also directly influence the marketing strategic factors, which enhances the feature compatibility ratio for the e-commerce portals users. In this paper, the unique collaborative ranking model based upon the k-nearest neighbour has been proposed along with the content-based PAT feature algorithm. The proposed model has been uniquely defined to add the higher flexibility and to produce the highly compatible lists over the e-commerce portals for the generation of better results. The proposed model has been evaluated under the time-domain and for the uniqueness. In both of the domains the proposed model has outperformed the existing model based upon neural network.

Keywords— E-book recommendations, library recommendations, digital library ranking engine, semantic recommendation system.

I. INTRODUCTION

The aim of ranking systems is to assist users in finding their way through huge databases and catalogues, by filtering and suggesting relevant items taking into account or inferring the user's preferences (i.e., tastes, interests, or priorities).Figure 1. below depicts the basic input requirement to build up a recommendation system to that it will produce recommendation list based on the analysis of that input. Product ranking is the way of classifying products on the product page that are shown to the users on various e-commerce shopping portals. It provide users with ease of choosing, viewing or purchasing various products according to their popularity, no. of views etcetera. Today, majority of people worldwide like to shop online; this is why today online shopping portals are becoming a booming business. As there are variety of products displayed on the website it is difficult for the user to view hundreds of thousands of products at the same time.

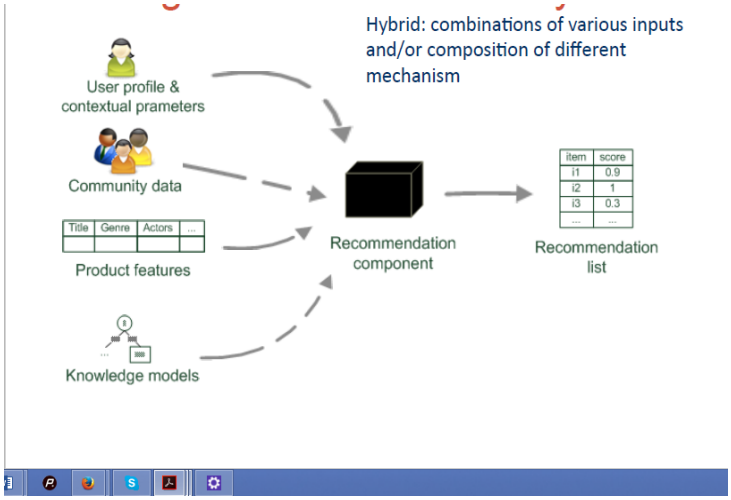

Figure 1. Process of recommendation system

However, to deal with such problem, the proposed system, rank the products usually with three factors i.e. popularity, accessibility and trust (PAT). The proposed system decides which products in the online e-commerce portals is the most similar one with the query product. To speed up the ranking system, use the grid-based indexing to obtain the nearest neighbours of the product listing, exact matches are retrieved. The products related to the query information are collected, ranked and saved for the various ranking requirements. Indexing can be performed in vector space to improve ranking system execution speed

II. RELATED WORK

In [5] authors have proposed a recommender system using Fuzzy Linguistic approach that has been tested in a University Digital Library. Their main focus was on quality factor of the items and their relevance to provide users at ease to access the appropriate research resources. In [6] authors have proposed a recommender algorithm and compare it with three common techniques to solve the recommendation issue namely collaborative filtering, cluster model and search based method using large datasets. Hence, their algorithm produced real-time, high quality recommendations to the users. IN [3] authors have worked on the E-commerce Website Ranking Using Semantic Web Mining and Neural Computing. The authors have discussed the design of the new semantic web based E-commerce page ranking model based on neural network. The website ranking tool has been realized by using the various relevant features to rank the web in the real time in order to facilitate the customers by providing the rank. Also the recommender system can be utilized for the recommendation on the various products displayed on the various website portals. In [1] authors have given a comprehensive overview on the advanced conceptual structures of schema.org for e-commerce, including patterns for ownership and demand, and have present the full tool chain for producing and consuming respective data. Also the authors have explained the long-term vision of Linked Open Commerce, and discussed advanced topics, like access control, identity and authentication, micropayment services, and data management issues from the publisher and consumer perspective. The authors have also covered research opportunities resulting from the growing adoption and the respective amount of data in RDFa, Microdata, and JSON-LD syntaxes. In [7] authors presented an article in which they have discussed about the various queries made by users on webpages, though they defined a class namely Package Query and a frequent subclass Skyline package queries using semantic web to give the quality results to the users. In [9] author proposed Mini-ME , a mobile inference engine designed from SwoT that supports semantic web technologies and provide performance evaluations both on a smartphones and PC's.

III. EXPERIMENTAL DESIGN

The proposed model is based upon the e-commerce product ranking model. The producing ranking models are required in the multivariate applications over the online e-commerce/shopping portals. The efficient product listing requires a number of computations to prepare the accurate product ranking. The accuracy in the product ranking model is not-measureable in the exact terms. Hence the performance evaluation of the product ranking models is evaluated from the resource usage and the entropy. The proposed model utilizes the query based ranking and the user-profile based ranking, which makes the proposed model a hybrid ranking model. The hybrid ranking model primarily utilizes the dual class classification, where the popularity, accessibility and trust based factor for the modelling of the product ranking model. The proposed model utilizes the k-nearest neighbour classification algorithm for the arrangement of the collaborative ranking based upon the user profile.

The existing model has been improved by using the hybrid approach for the purpose of recreation of the robust ranking model. The local and global parameters has been utilized using the PAT (popularity, accessibility and trust) based factors. The k-nearest neighbour algorithm has been utilized for the classification of the users under the collaborative approach. The proposed collaborative approach has been utilized for the user similarity matrix for the purpose of finding the neighbour entity relationship. The following algorithmic steps have been followed to achieve the k-NN classification goal over the PAT features:

Algorithm 1: PAT based k-NN classification for the user similarity evaluation

1. Restructure the PAT based data ranking list
2. Pass the evaluation rate to the k-NN method
3. Get the size in number of rows of the ranking list
4. If matrix does not match the input sequence
    a. Return the program with the collaborative ranking
5. Acquire the user history information from the web access log database
6. If the user history is empty
    a. Return the program with the collaborative ranking
7. Acquire the user training data (UTD)
8. Load the historical data into the runtime memory
9. Perform the averaging factor over the historical data
    a. Return the virtual user profile vector (VUPV)
10. Start the iteration for similar evaluation
11. Acquire the averaged PAT feature for the current row
12. Acquire the virtual user profile vector
13. Evaluate the individual distance between individual components
    a. Return the similarity matrix
14. Compute the cumulative phase distance between the VUPV and UTD
15. Update the kNN similarity matrix
    a. Return the similarity matrix

16. Obtain the most similar candidates from the similarity vector
17. Perform the averaging factor to prepare the representative matrix (RM)
18. Re-compute the content based ranking list against the RM
19. Restructure the content based ranking list
20. Return the final product ranking matrix

IV. RESULT ANALYSIS

The proposed model named PAT model, which has been compared against the existing model of product ranking based upon web parameters over the given e-commerce product list in this simulation scenario with similar structure and environment. The detailed analysis of the simulation results has been conducted by analysing the results obtained from the existing and proposed model simulations. The proposed model of proposed model has been described efficient and effective while evaluated on the basis of the projected resources and entropy for the coalition of the network load and uniqueness respectively. Minimum of the 10 percent improvement has been observed in the favour of the proposed model when compared on the basis of various performance parameters in the variety of simulation scenarios. The adverse situations have been analysed theoretically and the effective in the results has been observed in the proposed model to mitigate the errors and repetitions with the new product ranking model than the existing model. The newly proposed scheme has been primarily evaluated for the utilization of the resources while using the ranking model the given product data in the simulation. The utilization of the resources, measured in the form of projected resources has been recorded significantly lower than existing model, which clearly indicates the robustness of the proposed model. The following table (Table 1) indicates the robust performance of the proposed model.

Table 1: Projected Resources comparison between existing and proposed model

Projected Resources	**S1:E100**		**S2:E500**		**S3:E1000**	
	Existing	**Proposed**	**Existing**	**Proposed**	**Existing**	**Proposed**
Average	2.6499	1.5469	13.3108	7.7344	26.4594	15.4688
Minimum	1.8221	1.0977	7.8656	5.4883	15.0027	11.2500
Maximum	3.7667	1.9648	19.7488	9.8242	38.6813	19.6484

Projected Resources	**S4:E2000**		**S5:E5000**	
	Existing	**Proposed**	**Existing**	**Proposed**
Average	52.7602	30.9375	110.9954	77.3438
Minimum	35.7773	21.9531	80.7805	54.8828
Maximum	77.2214	39.2969	160.2389	98.2422

Additionally, the entropy parameter signifies the higher uniqueness level in the ranking data using the proposed model than the existing model. The proposed model evaluation over the entropy has been described in the following table (Table 2). The high uniqueness of the product table data increases the additional robustness to the ranking models associated with the arrangement of the entries in the perfect form over the given set of data. The result of the proposed model has been observed nearly double than the existing model, which can be clearly indicated by the following table:

Table 2: Entropy comparison between existing and proposed models

ENTROPY	**S1:E100**		**S2:E500**		**S3:E1000**	
	Existing	**Proposed**	**Existing**	**Proposed**	**Existing**	**Proposed**
Average	2.0207	2.4678	2.0131	2.6344	2.0712	2.7662
Minimum	1.7632	2.0544	1.6090	2.1550	1.7707	1.9910
Maximum	2.2706	2.9963	2.4393	3.1638	2.3884	3.2897

ENTROPY	**S4:E2000**		**S5:E5000**	
	Existing	**Proposed**	**Existing**	**Proposed**
Average	1.9829	2.5944	2.1466	2.5038
Minimum	1.5884	1.9055	1.9260	1.9444
Maximum	2.3225	3.0931	2.2644	3.1026

The result of the proposed model has been observed nearly double than the existing model, which can be clearly indicated in the above table.

V. CONCLUSION

The proposed model is based upon the k-nearest neighbour classification based collaborative approach for the preparation of the product ranking over the input e-commerce product listings. The proposed model has been developed by aiming at the robust and flexible listings to facilitate the users for the easy navigation towards the most related, most popular and similar interest based product listings. By using these practices, the e-commerce engines can also enhance the productivity by increasing the overall sales. The proposed model aims at the popularity based factors, while keeping the product page trust level as well as the accessibility for the final listings of the product data. The proposed model has been evaluated for the projected resource, overall time complexity and the entropy to measure the uniqueness of the entries in the product data. The experimental results have shown that the proposed model effectively outperformed the existing model. In the future, the multi-level lists can be reproduced in order to create the cumulative list for the mobile users, as the mobile users have smaller screen where the short product list can be shown only. Also the proposed model can be enhanced by combining the random forest, Coforest or support vector machine based probabilistic classification model with the non-probabilistic k-NN model for the realization of the robust ranking model.

ACKNOWLEDGMENT

Nothing concrete can be achieved without an optimum combination of inspiration and perspiration. Like all the others studies, this work are also the result of the interaction of number of minds that directly contributed to completion of this paper. First and foremost I like to thank God, the Almighty for his showers of blessings. I would also like to express a deep sense of gratitude and thanks profusely to my guide Mrs. Pravneet Kaur, Assistant Professor, Computer Science & Engineering Department, CGCTC, Jhanjeri. Without her wise counsel and able guidance, it would have been impossible to complete the work in this manner. Last but not the least; I am highly thankful to my friends and family for helping me round the clock, which has been a source of great encouragement to me.

REFERENCES

[1] Hepp, Martin. "The Web of Data for Online Libraries: Schema. org and Good Relations for Researchers and Practitioners." In *Engineering the Web in the Big Data Era*, pp. 723-727.Springer International Publishing, 2015.

[2] Furukawa, Takao, Kaoru Mori, KazumaArino, Kazuhiro Hayashi, and Nobuyuki Shirakawa. "Identifying the evolutionary process of emerging technologies: A chronological network analysis of World Wide Web conference sessions."*Technological Forecasting and Social Change* 91 (2015): 280-294

[3] Verma, Neha, Dheeraj Malhotra, Monica Malhotra, and Jatinder Singh. "Online Libraries Website Recommendation Using Semantic Web Mining and Neural Computing." *Procedia Computer Science* 45 (2015): pp. 42-51, ELSEVIER.

[4] Greg Linden, Brent Smith, and Jeremy York "Amazon.com Recommendations Item-to-Item Collaborative Filtering" Published by the IEEE Computer Society.

[5] Tejeda-Lorente, Álvaro, Carlos Porcel, Eduardo Peis, Rosa Sanz, and Enrique Herrera-Viedma. "A quality based recommender system to disseminate information in a university digital library." *Information Sciences*261 (2014): 52-69.

[6] Greg Linden, Brent Smith, and Jeremy York "Amazon.com Recommendations Item-to-Item Collaborative Filtering" Published by the IEEE Computer Society

[7] Sessoms, Matthew, and Kemafor Anyanwu. "Enabling a Package Query Paradigm on the Semantic Web: Model and Algorithms."In *Transactions on Large-Scale Data-and Knowledge-Centered Systems XIII*, pp. 1-32.Springer Berlin Heidelberg, 2014.

[8] Malhotra, Dhairya. "Intelligent web mining to ameliorate Web Page Rank using Back-Propagation neural network." In *Confluence The Next Generation Information Technology Summit (Confluence), 2014 5th International Conference*-, pp. 77-81. IEEE, 2014.

[9] Scioscia, Floriano, Michele Ruta, Giuseppe Loseto, Filippo Gramegna, SaverioIeva, Agnese Pinto, and Eugenio Di Sciascio. "A Mobile Matchmaker for the Ubiquitous Semantic Web." *International Journal on Semantic Web and Information Systems (IJSWIS)* 10, no. 4 (2014): 77-100.

[10] Mital, Monika, AshisPani, and Ram Ramesh. "Determinants of choice of semantic web based Software as a Service: An integrative framework in the context of e-procurement and ERP." *Computers in Industry* 65, no. 5 (2014): 821-827.

[11] D. Jannach, M. Zanker, A.Felfernig and G. Friedrich, Recommender Systems – an introduction Cambridge University Press, 2010

[12] D.T. Green and J. M. Pearson, "The examination of two web site usability instruments for use in B2C Online Libraries organizations," *Journal of Computer Information Systems*, Vol. 49, No. 4, 2009, pp. 19-32

[13] T. Wang and Y. Lin, "Accurately predicting the success of B2B ecommerce in small and medium enterprises," *Expert Systems with Applications*, Vol. 36, No. 2, published by Elsevier, 2009, pp. 2750–2758.

