Personalize Recommendation Approach for Web Search in E-Learning

S.M. Momin[1], A.B. Rajmane[2]

[1,2]Department of Computer Science and Engineering, Ashokrao Mane Group of Institutions, Vathar, India

---

Abstract- Nowadays, new technologies and the fast increase of the Internet have made access to information easier for all kinds of people, building new challenges for education when utilizing the Internet as a tool. E-Learning provide students with choices and initiative, however, results in much challenge in matching the needs of students with different backgrounds and learning preferences due to information overload. Facing diverse learning resources, students have difficulties in making appropriate choices to meet their learning objectives. One of the best examples is how to personalize an E-Learning system according to the learner's requirements and knowledge level in a learning process. This system should adapt the learning experience, according to the goals of the individual learner. In this paper, we present a recommender E-Learning approach which utilizes recommendation techniques for educational data mining specifically for identifying E-Learners' learning preferences. E-Learning recommendation system helps learners to make choices without sufficient personal experience of the alternatives, and it is considerably requisite in this information explosion age. In our study, the user-based collaborative filtering method is chosen as the primary recommendation algorithm, combined with online education. We analyze the requirement of a web based E-Learning recommendation system.

The proposed system is based on four modules, namely Web search Module, student Profiling Module, Behavioral Activity analyzer module, recommendation module. The web search Module is a way of searching anything that user or student wants from Google search engine. A student Profiling Module takes Students all Personal and Academic Information, Behavioral Activity analyzer module is used to identify learners learning preferences and all activities which are done at the time of web surfing by students and a recommendation module which pre-processes data to create a suitable recommendation list and predicting the student interest domain. After recommendation process we calculate the Knowledge Point (KP) of particular student based on KP value it categories the student into three levels 1. Beginner 2. Intermediate 3. Master. Several techniques such as classification, clustering and association rules are used to improve personalization with filtering techniques to provide a recommendation and assist learners to improve their performance.

Keywords - E-Learning, recommender system, collaborative filtering, Student profiling, Classification, Weight Ranking, Knowledge Point (KP).

1. INTRODUCTION

Personalized recommendation initially originated from the electronic commerce, which is an effective way to reduce the information overload. It can provide personalized information and products to assist consumers in making decisions in terms of consumers' interests and preferences [1]. Technology Enhanced learning is the application of information and communication technologies for teaching and learning. Collaboration based Filtering is proposed in this paper to investigate the personalized recommendation in E-Learning.

Recommendation Systems (RS) are software tools based on machine learning and information retrieval techniques [2] that provide recommendations for potential useful items to someone's interest. Most of the modern E-Learning systems are still producing the same educational resources in the same way to learners with various profiles [3]. In general, to enable personalization, existing systems use one or more types of knowledge (learning process knowledge, learners' knowledge, learning materials knowledge, etc.) and personalization in E-Learning systems involve adaptive course delivery, adaptive collaboration support, adaptive interaction and content discovery [12]. Due to a large amount of learning resources on the web, it is difficult to find learning resources associated to learner request [4].

E-Learning recommender systems intend to recommend a sequence of items to learners, that is, to recommend the most efficient or effective paths within large learning resources to achieve a specific competence. This paper presents a recommender system for e-Leaning personalization based on learners learning activities and performance. It means personalization approach of giving learning resources for active learners in the E-Learning system. This system recommends some learning resources based on learner's profile, level of knowledge, and other learner's activities. Also, the system provides the ability to track learner achievement based on practical tests and exercises and observe the learner's performance in order to supervise and support the learners, now a day's commercial search engines often place sponsored advertisements in the form of CPM (Cost per thousand viewers), CTR (Click-through rates), CPC (Cost per click), or CPA (Cost per action) [7] over relevant items which also distract students from choosing the right sources of content from the returned search results, because of this the student will not get the required knowledge and it take the unwanted links because of this it get confuse.

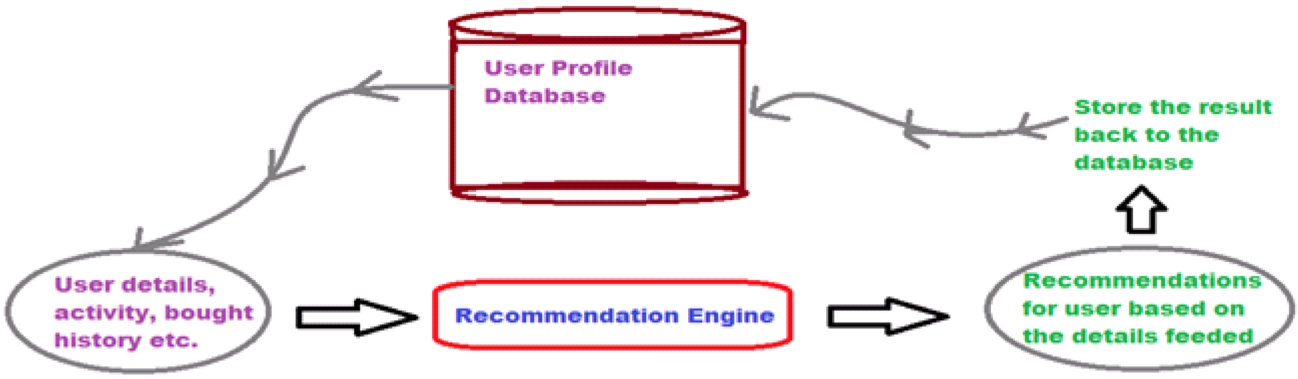

Figure 1: General Flow of Recommendation System

Figure 1: General Flow of Recommendation System

2. RECOMMENDATION SYSTEM -

This work is based on a recommendation system, it is a subclass of the information filtering system that seeks to predict the "rating" or "preference" a user would give to an item. Recommender system envelops a class of methods and calculations which can propose "relevant" things to clients. In a perfect world, the proposed things are as pertinent to the client as would be engaged, so the client can draw in with those things: YouTube recordings, news stories, online items, etc. [9].

Items are ranked according to their relevancy, and the most relevant ones are shown to the user. The relevancy is something that the recommender system must determine and is mainly based on historical data. If you've recently watched YouTube videos about elephants, then YouTube is going to start showing you a lot of elephant videos with similar titles and themes [14].

Recommender systems are generally divided into two main categories: 1. Collaborative filtering 2. Content-based systems.

2.1 CLASSIFICATION OF RECOMMENDATION SYSTEMS IN

E-LEARNING-

This section presents a classification of Recommendation systems in E-Learning taking into account several factors. First, Subsection reviews the preprocessing techniques, features and models used in order to represent the search results. Next, Subsection presents the most popular clustering algorithms employed to group the web pages related to each individual with the same name. Following that, the main part of Recommendation system that take into account the actual recommended links which shows as a result respective a search [9].

The phase of Recommendation systems in E-Learning systems consists of representing the search results in order to treat them automatically in a correct way. This phase is usually composed of the following steps:

Preprocessing:

In this step, the search results are processed from their original format by means of several techniques.

Feature selection:

The main goal of this step is to select suitable features to distinguish different individuals with the same name correctly.

Representation model:

In this step, the search results are represented by means of a certain recommended links. On the one hand, these models assign a value to each feature representing its importance with respect to the web page it belongs to, or to all the search results retrieved by the search engine. On the other hand, these models allow the comparison of web pages by means of behavioral operations.

2.2 COMPARISON OF CLUSTERING ALGORITHMS -

Learning Method	Loss Function	Number of clusters: Predetermined or Data-dependent	Cluster shape: isotropic or anisotropic?	Parameter Estimation Algorithm
K-means	Within-class squared distance from mean	Predetermined	Isotropic	K-means
Gaussian Mixture Models (identity covariance)	$-\log P(X)$, (equivalent to within-class squared distance from mean)	Predetermined	Isotropic	Expectation Maximization (EM)
Single-Link Hierarchical Clustering	Maximum distance between a point and its nearest neighbor within a cluster	Data-dependent	Anisotropic	Greedy agglomerative clustering
Spectral Clustering	Balanced cut	Predetermined	Anisotropic	Run Laplacian Eigenmaps followed by K-means or thresholding eigenvector signs

Table 1: Comparison of Clustering Algorithms

3. LITERATURE REVIEW-

Application Research on Personalized Recommendation in distance education [1]

Personalized recommendation in distance education, it represents multiple recommendations algorithms which say the latest achievements of research on personalized recommendation service. In this research, the user's interest information and user modeling are getting collected, item matching have made up the core content of recommendation system [1].

The limitations of this survey are, it is less accurate and inefficient in finding user's interests in personalized recommendation. It works on how often and how to update the user model so that it can provide accurate recommendation results to users. This research fails to express the basic resource in a standard method and improve the shared degree of recommendation on an open system.

E-commerce Personalized Recommendation System Based on Multi Agent [2]

Multi-Agent to E-commerce personalized Recommender System, and design E-commerce personalized Recommender System based on MAPRS (Multi-Agent Personalized Recommendation System). Off-line recommendation and on-line hybrid recommendation are used to construct the core recommender model under the intelligent control. This research presents the function and design ideas of various components of the system.

This survey has limitations of Security and User profile classification problem. This may require dealing with the user's privacy. The creation of the full user profile is difficult

A Graph-Based Taxonomy of Recommendation Algorithms and Systems in LBSNs [3]

The exploitation of geographic hierarchy is mostly seems to exploit geographic hierarchy information. That is, there is a universal system which defines a geographic hierarchy. This research is to seek for recommendation algorithms in LBSN (Location-based Social Networks) [3], which are able to provide more accurate and justifiable recommendations. Moreover, during the past decade, many different websites and many algorithms were introduced to provide suggestions close to user needs. This survey presented 43 recommendation algorithms in LBSNs and compared 16 real-life LBSNs,

This Survey have a limitation of, exploration of geographic hierarchy it have to consider the location factor and time factor correctly, if the location and time factor values are mismatch then the result of recommendation is fails [3].

Personalized Image Recommendation for Web Search Engine Users [4]

This research is developed for personalized search system is being proposed alternate query generator is used to capture all the senses of the main query and assist the user with alternate queries [2]. It proposes personalization based profile, click history and last action performed by the user is used to improve the ranking of search results. It proposed personalize system architecture in two layers: 1) data presentation layer 2) Data collection layer.

The limitation of this Survey is it creates the language gap between the user and search engine.

3.1 Comparison of different web search engines:

Sr. No.	Name of Search Engine	Review	Personalize Recommendation	Search in Entire Web	Dynamic Profiling
1	Google scholar	Powered by Google, this search makes search simple	No	Limited	No
2	Search Guide	It Tool is used to search , it exhibits trails of research from the past search of user	No	Yes	No
3	LOBSTER	Specialized search assistant Tool based on the Google search, compare some advance features that help the user to find the correct findings	No	Yes	No
4	iSEEK	Created particular for the understudies or students it is non-commercialized and it present the result in editor view	No	Limited	No

Table 2: Comparison of different web search engines for educational purpose.

3.2 The findings notes from literature Data:

1) Web search engines provide limited capabilities in personalizing the search results according to the student's profile even though the Web search engines were the most popular tool used by students in searching for educational materials.

2) The majority of research looking at adaptive learning approaches had concentrated mainly on the personalization of learning materials in various E-Learning systems but not on the Web search engines.

3) Several studies have examined the use of the Web search engines as E-Learning tools but thus far, from the best of my knowledge no study was performed in providing personalized recommendations to students using Web search engines.

4) Groupization algorithm for distributing the Students into the several categories like Master, Intermediate, and Beginner has not yet been implemented in personalizing the delivery of learning materials through Web search engines.

4. DEVELOPMENT OF A PERSONALISED WEB SEARCH RECOMMENDATION SYSTEM FOR

E-LEARNING-

I. WEB SEARCH MODULE-

The web search module provides user Interface from where the student or user can search anything on the web, before the user goes for the web search we take some information about a user while registering. At the time of user creation, we take some personal information of the user like user enrollment number, department, year of studying and the user will set the password manually [11].

After taking some general information from the student, we allow user to login, when the login is done students have to fill the registration form in that student general information like his hobbies, technical interest, date of birth and the academic information.

With the help of academic information, we calculate the Knowledge Point. After that, the user will allow to search anything that he wants. All user search records will be maintained in the database. Simply all search histories are stored into the database [19].

In the database, we stored the link which is visited by the user or student, how many times it scrolls the page, how many times it click on the page and which link it click, how much time it spent on the particular site, like that all information is stored on the database.

This module intends to build and keep up the profile of every individual student. It permits a system to comprehend distinctive adapting needs and abilities of every individual student. It then uses this information to improve the consistency of the returned Web search results by selecting the most related personalized recommended links. Basically, this module will give the facility to log in to the user, it provides the authentication to the student, in that we get student academic information based on this we calculate the KP of the particular student [15].

II. STUDENT PROFILE CLASSIFICATION-

The student profiles are classified according to their academic performance and learning behaviour based on this we calculate the KP (Knowledge Point), based on the KP we divide the individual students in the following categories that are - a) Master b) Intermediate c) Beginner

To achieve this, the profile classification module has two functional components:

A) **Academic Record Analyzer**

B) **Behavioural Activity Analyzer**

A) **Academic Record Analyzer-**

The Academic Record Analyzer module is responsible for recognizing the login student by recovering their profiles from the Student MIS. It primarily recovers understudy's past and current scholastic data [5].

B) Behavioral Activity Analyzer-

The Behavioral Activity Analyzer module is liable for constantly observing and catching the student learning behaviour. This is accomplished through their Web search exercises. The data caught involves the occasions they sign in, the quantity or number of searches per login, the issued queries, the case documents, page names, page sizes, average scrolls, links clicked, and time spent [5] [8].

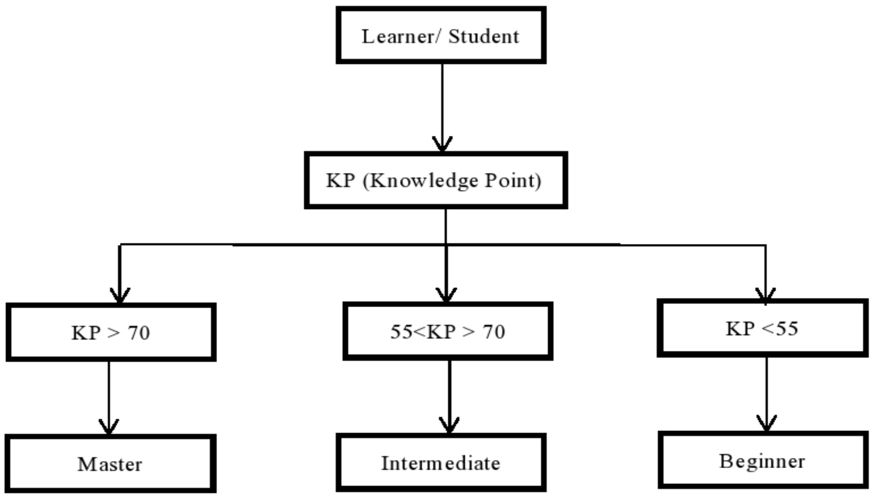

Figure 2: Decision tree for student profile classification.

1) Students with KP higher than 70 will be assigned a Master Class.

2) Students with KP are Between 55 to 70 will be assigned an Intermediate Class.

3) Students with KP less than 55 will be assigned a Beginners Class.

III. CONTENT RE-RANKING MODULE-

The fundamental reason for content re-ranking is to return query items into a custom sequence and organize them in a manner that is progressively relevant to the gathering individuals dependent on part comparability and level of inclination. That can be accomplished with the assistance of Groupization Technique.

Web search engine having personalized highlight check returns a different search result for unique people. Alternatively, it can coordinate the results in an alternate manner as per the client's purpose [15], dependent on the client's data needs and inclinations. However, it is trying to gather client's information which enough to comprehend the client's exact needs and inclinations. One approach to address this difficult task is by consolidating related information assembled from others with comparative profiles when building personalization features. In the 'Groupization' Technique, personalization is set up by allotting higher loads to pages that are suitable for the transcendence of the individuals from the gathering. This is accomplished by joining with each member report, term frequency and Web histories [17].

The requirement for recommendations has emerged in many situations, for example, recommending a goal of the movement for a family to spend an occasional break, a film for a gathering of companions to observe together, a great restaurant for colleagues to have a weekday lunch or even a lot of learning materials to help students of similar profiles in understanding a point. Naturally, things that are interesting to an individual from a group can likewise be interesting to different individuals from a similar group. Besides, students learn better in all around organized, helpful conditions instead of in the generally organized study hall condition.

Groupization is one of the strategies that can improve the estimation of community search devices. This can be performed by utilizing personalization with shared interests. In a collaborative filtering technique, the recommender system proposes things depend upon how identical clients want the thing. It closures to reveal there with comparative or related concerns/interests [18]. However, the matches are developed depending on queries on recommendations or suggestions or network connectedness rankings. Groupization considers the common interests of the gathering individuals in order to create better personalization results (positioning level). The common data is normally insignificant or inaccessible to web crawlers which encompass data [10].

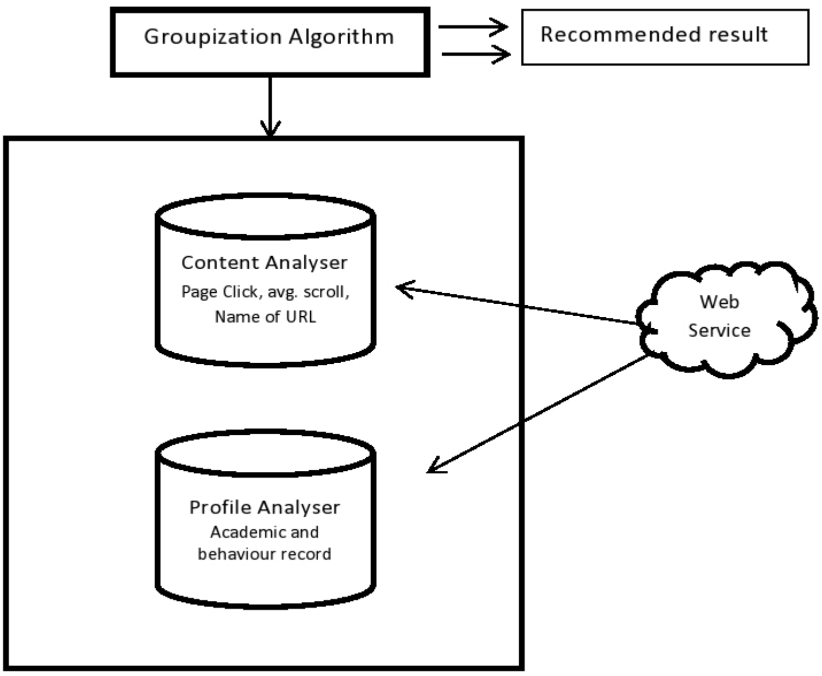

Figure 3: Content Re-ranking

IV. RECOMMENDATION MODULE-

In this module we get the recommended results, this can be achieved with the student behavior, in that we consider the Academic Record Analyzer, Behavioral Activity Analyzer [5] [8].
In that, we are making use of Collaborative filtering algorithm, that will helpful to recommend the links.

That can be achieved by the following steps

1. Find the top similar users with respect to the particular user.
2. Predict the users rating on an item based on other users.
3. Recommend the items which have higher predicted value

The **Collaborative filtering algorithm is used** *Centered Cosine technique.*

Centered Cosine can be implemented by,

* We take the mean of the user ratings and subtract that means from all individual ratings divided by the total number of ratings by users.

* For all the links, where there is no rating by users, we replace it with 0.

Basically, this module will show a recommended link extracted from the Google, on the basis of the student academic behavior and internet behavior. The main objective of this module is to show the appropriate result to the user.
The first step in creating a recommendation engine is gathering data. Data can be either explicit or implicit. Explicit data consists of data inputted by users, such as the interest of the user. Implicit data might include page visited history, page views, clicks, and search log, time spend, and this data is collected for every user who visits any given site [14].

Behavior data is easy to collect because you can keep a log of user activities on your site. Collecting this data is also straightforward because it doesn't require any extra action from the user; they're already using the application, after all. The downside of this approach is that it's harder to analyze the data. For example, filtering the necessary logs from the less important logs can be cumbersome [13].

The next step is to filter the data to get the relevant data necessary to provide recommendations to the user. We have to choose an algorithm that would better suit the recommendation engine from the list of algorithms explained above. Some types of filters are:

- **Content-based**: A popular, recommended product has similar characteristics to what a user views or likes.

- **Cluster**: Recommended products go well together, no matter what other users have done.

- **Collaborative**: Other users, who like the same products as another user views or likes, will also like a recommended product.

Collaborative filtering enables you to make product attributes theoretical and make predictions based on user tastes. The output of this filtering is based on the assumption that two users who liked the same products in the past will probably like the same ones now or in the future [11].

5. EVALUATION METHOD

To estimate the ease of use, usefulness, and effectiveness of the proposed personalized Web search recommendation system, experimentation was designed and conducted for the Diploma students of our College. A prototype was developed using PHP and MySQL and used on a local machine. Google API was used to integrate the Google search engine with the prototype. For the purpose of the experiments, the prototype was made available for the participants to access the system by creating individual accounts

5.1. PARTICIPANTS

In total, 30 students were approached to participate in the experiment. The second-year students from the Faculty of Computer Engineering were chosen as participants. The intention of working with Computer Science students was derived from the assumption that they were more technically stronger than others department students in using a search engine or any web activity. The experiment aims to measure how easy these students can search for their learning materials from the Google search engine by using our proposed system. Among the 30 invited students, all 30 students participated in the experiment. Three groups were made by considering students regardless their academic records and web records.

5.2. PROCEDURE

The experiments were conducted in random sessions. Each random student asked to use or surfing the internet with our giving prototype. The single restraint placed on the participants throughout the session was to search for the solutions by employing the Google search engine that is accessible from the proposed system. The browsing history, session logs, and time spent on the internet were recorded. Based on the interest and the response of the user, the recommended link is shown into the user interface, the user or the student will easily search the anything they want and it the result will sow into the two planes first is the normal Google results and second is the recommended result based on their past record and interest, the recommended result will show the top K most links which are highly accessed by the other user, in that framework we are using the Page Ranking algorithm to calculate the weight of the particular accessed page. There are two phases of data collection. First is the students' personal and academic information which was automatically retrieved from the Students MIS and stored in a local server. Second is the students' learning activities which were obtained from their browsing histories and session data were collected while they were using the proposed system.

6. RESULTS AND DISCUSSION

ANALYSIS OF DEVELOPMENT WITH GRAPH-

To recall, the vastness and diversity of the contents in the Web has efficient to the need to revisit the concept of one search results to all user's proposition of the current Web search engine. Even though there are indeed personalized search applications to support users with their studying, shopping, travelling, entertainment or personal requirements, we believe that one community that also wishes to urgent consideration is the students where recommendations for suitable E-Learning materials from the Web need to be addressed. To match this gap, we designed and developed a novel personalized recommendation system for the Web search engine users. It aims to deliver a more personalized search result recommendation that matches individual student profiles. In this study, the user level of satisfaction in terms of the ease of use and usefulness of the system was measured. We also evaluated the searching performance and time taken by the users of our students while using the proposed system. The search results accuracy such as the precision and recall evaluation.

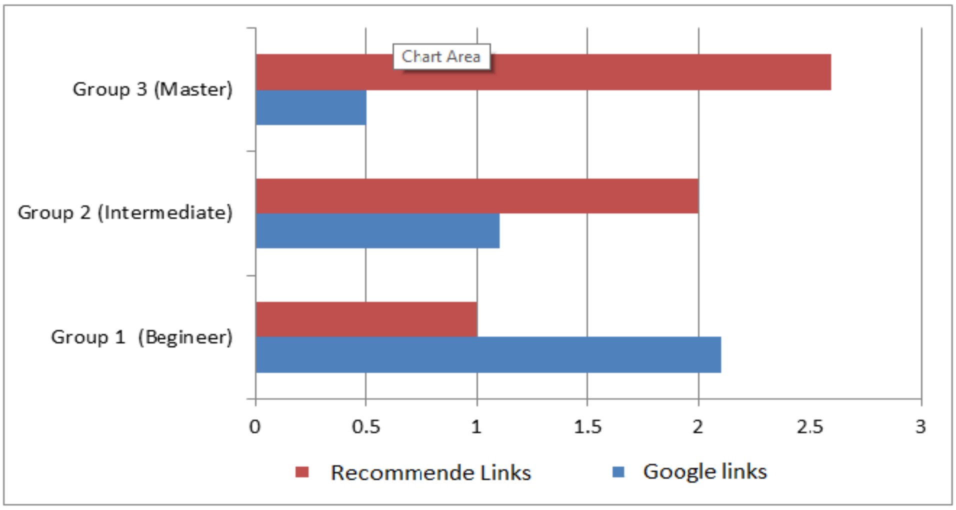

Figure 4: Google and Recommended Link Clicked by Students

In order to determine the students' preference towards the links displayed in each search result, we monitored their link selections. Figure 4 shows the total number of clicks on the Google returned links and the recommended links of each experimental group. When comparing it was observed that the number of clicks identified on the recommendation links increases while the number of clicks on Google returned links decreases. This exhibit an inverse relationship between recommendation links and the Google returned links when compared sequentially among the groups. One possible reason for the high number of clicks on the links displayed by Google is due to the students' familiarity with Google as compared to the recommended links section. The exciting finding obtained from this experiment shows that there is an incremental interest to select the recommended links. As mentioned in the earlier, as the participations of each group increases, their browsing histories also become richer hence making it possible for the proposed system to recommend links that were personalized to their profiles.

(b)

Figure 5: Search Interface and Recommendations List Display for (a) Google Default Result (b) Recommended Result

Note that the Figure 5 (a) displays the typical search results returned by Google search engine which are the same for all students. On the other hand, the recommended links returned by the system (as shown in Figure 5 (b)), appears to be dissimilar for different types of student profiles. Each profile belongs to one of the following three groups: Beginner, Intermediate, and Master.In above, the student search about the network setting. At the time of entering the query the Google search engine shows the some result or links that are common for all, In our recommended search on the search window one link is appear that is "for recommended result click here ", while user clicking on that link the user will get the recommended result with page weightage as per the search.

7. DISCUSSION AND CONCLUSION

Recommendation systems give the opportunity to learners to lead a successful learning experience by getting the right learning materials that suit their needs at the right time. The results from the experimental study revealed that it gained a significantly better learning comprehension as compared to the other. This finding implies that the personalized group-based recommendation system is helpful in improving the effectiveness of the students Web search and learning performance. The learning recommendation is designed to help student achieving the learning outcomes of the materials. The assessments of the student understanding suggest that the recommendation system is capable to increase the student score significantly [19].

The Main limit of previous system is the Google search engine provides only 100 free searches in a day that limitation is removed from this framework, the user can search unlimited search freely. To provide a personalized Web search, the recommender system needs to access users' personal profiles, description of user interest and user behaviour [17]. However, it is challenging for any recommender system to collect user data [18], thus more innovative approaches need to be considered.

In this study, the proposed system employed in the educational institutional environment, where students data can be obtained from the Students Management Information System.

Future work should consider that the possibility of incorporating student's social identities from their social networks for richer student's profiles and more desirable recommendation outcomes would also be an appealing topic to explore.

REFERENCES-

1. Jiang, Y., Zhan, H., & Zhuang, Q. (2010). Application research on personalized recommendation in distance education.
2. Guo, Y., & Liu, Q. (2010). E-commerce personalized recommendation system based on Multi-Agent
3. Kefalas, P., Symeonidis, P., & Manolopoulos, Y. (2016). A Graph-Based Taxonomy of Recommendation Algorithms and Systems in LBSNs. IEEE Transactions on Knowledge and Data Engineering.
4. K. D. Rajab, "the effectiveness and potential of E-Learning in war zones: An empirical comparison of face-to-face and online education in Saudi Arabia," IEEE Access, vol. 6, pp. 6783–6794, 2018.
5. Y. Li, J. Luo, and T. Mei, "Personalized image recommendation for Web search engine users," in Proc. IEEE Int. Conf. Multimedia Expo, Jul. 2014, pp. 1–6.
6. H. Wang and K. Wong, "Recommendation-assisted personal Web," in Proc. IEEE 9th World Congr. Services, SERVICES, Jun. 2013, pp. 136–140.
7. Pessemier, T. D., Deryckere, T., & Martens, L. (2010). Extending the Bayesian Classifier to a Context-Aware Recommender System for Mobile Devices. 2010 Fifth International Conference on Internet and Web Applications and Services.
8. C. Yathongchai, T. Angskun, W. Yathongchai, and J. Angskun, "Learner Classification based on learning behavior and performance," in Proc. IEEE Conf. Open Syst., Dec. 2013, pp. 66–70.
9. Li, H., Wang, L., Du, X., & Zhang, M. (2017). Research on the Strategy of E-Learning Resources Recommendation Based on Learning Context. 2017 International Conference of Educational Innovation through Technology
10. Using Collaborative Filtering Algorithms as eLearning Tools. (2009). 2009 42nd Hawaii International Conference on System Sciences.
11. Wang, G. T., Xie, F., Tsunoda, F., Maezawa, H., & Onoma, A. K. (n.d.). Web search with personalization and knowledge. Fourth International Symposium on Multimedia Software Engineering
12. T.-C. Yang, G.-J. Hwang, and S. J.-H. Yang, "Development of an adaptive learning system with multiple perspectives based on students' learning styles and cognitive styles," Educ. Technol. Soc., vol. 16, no. 4, pp. 185–200, 2013.
13. K.W.-T. Leung, D. L. Lee, and W.-C. Lee, ``Personalized Web search with location preferences,'' in Proc. Int. Conf. Data Eng., 2010,
14. Web search with concept-based user proles,'' ACM Trans. Internet Technol., vol. 11, no. 4, pp. 1_29, 2012.
15. M. R. Morris, J. Teevan, and S. Bush, ``Enhancing collaborative Web search with personalization,'' in Proc. ACM Conf. Comput. Supported Cooperative Work (CSCW), p. 481, 2008.
16. S. Sendhilkumar, K. Selvakumar, and G. S. Mahalakshmi, ``Application of fuzzy logic for user classification in personalized Web search,'' Int. J.Cybern. Inf., vol. 3, no. 3, pp. 23_49, 2014.

17. Pessemier, T. D., Deryckere, T., & Martens, L. (2010). Extending the Bayesian Classifier to a Context-Aware Recommender System for Mobile Devices. 2010 Fifth International Conference on Internet and Web Applications and Services.
18. Mohamed Suffian, M. D., Fairul Rizal, F., Loo, F. A., Aman, N. F., & Bajuri, N. (2016). Software capability rating using system testing scores. 2016 IEEE Conference on Open Systems.
19. M. Speretta and S. Gauch, ``Personalized search based on user search histories,'' in Proc. IEEE/WIC/ACM Int. Conf. Web Intell., Sep. 2005,pp. 622_628