PERSONALISED WEB SEARCH FOR E-LEARNING USING GROUP-BASED RECOMMENDATION APPROACH

MOHAMMAD MUSTANEER RAHMAN

FACULTY OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

PERSONALISED WEB SEARCH FOR E-LEARNING USING GROUP-BASED RECOMMENDATION APPROACH

MOHAMMAD MUSTANEER RAHMAN

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF COMPUTER SCIENCE

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Mohammad Mustaneer Rahman

Matric No: WGA140061

Name of Degree: Master of Computer Science

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”): PERSONALISED WEB SEARCH FOR E-LEARNING USING GROUP-BASED RECOMMENDATION APPROACH

Field of Study: E-Learning (Computer Science)

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature Date:

Subscribed and solemnly declared before,

Witness’s Signature Date:

Name:

Designation:

UNIVERSITI MALAYA

PERAKUAN KEASLIAN PENULISAN

Nama: Mohammad Mustaneer Rahman

No. Matrik: WGA140061

Nama Ijazah: Master of Computer Science

Tajuk Kertas Projek/Laporan Penyelidikan/Disertasi/Tesis ("Hasil Kerja ini"): PERSONALISED WEB SEARCH FOR E-LEARNING USING GROUP-BASED RECOMMENDATION APPROACH

Bidang Penyelidikan: E-Learning (Computer Science)

Saya dengan sesungguhnya dan sebenarnya mengaku bahawa:

(1) Saya adalah satu-satunya pengarang/penulis Hasil Kerja ini;
(2) Hasil Kerja ini adalah asli;
(3) Apa-apa penggunaan mana-mana hasil kerja yang mengandungi hakcipta telah dilakukan secara urusan yang wajar dan bagi maksud yang dibenarkan dan apa-apa petikan, ekstrak, rujukan atau pengeluaran semula daripada atau kepada mana-mana hasil kerja yang mengandungi hakcipta telah dinyatakan dengan sejelasnya dan secukupnya dan satu pengiktirafan tajuk hasil kerja tersebut dan pengarang/penulisnya telah dilakukan di dalam Hasil Kerja ini;
(4) Saya tidak mempunyai apa-apa pengetahuan sebenar atau patut semunasabahnya tahu bahawa penghasilan Hasil Kerja ini melanggar suatu hakcipta hasil kerja yang lain;
(5) Saya dengan ini menyerahkan kesemua dan tiap-tiap hak yang terkandung di dalam hakcipta Hasil Kerja ini kepada Universiti Malaya ("UM") yang seterusnya mula dari sekarang adalah tuan punya kepada hakcipta di dalam Hasil Kerja ini dan apa-apa pengeluaran semula atau penggunaan dalam apa jua bentuk atau dengan apa juga cara sekalipun adalah dilarang tanpa terlebih dahulu mendapat kebenaran bertulis dari UM;
(6) Saya sedar sepenuhnya sekiranya dalam masa penghasilan Hasil Kerja ini saya telah melanggar suatu hakcipta hasil kerja yang lain sama ada dengan niat atau sebaliknya, saya boleh dikenakan tindakan undang-undang atau apa-apa tindakan lain sebagaimana yang diputuskan oleh UM.

Tandatangan Calon   Tarikh:

Diperbuat dan sesungguhnya diakui di hadapan,

Tandatangan Saksi   Tarikh:

Nama:

Jawat

PERSONALISED WEB SEARCH FOR E-LEARNING USING GROUP-BASED RECOMMENDATION APPROACH

ABSTRACT

The increasing dependency of students on the Web for learning is fuelled by the increasing availability and unprecedented growth of the Internet. Popular Web search engines in the market which depend on the right use of keywords in order to search the relevant learning materials do not take into account the learning proficiency of their users. Consequently, students will receive the same set of search results when the same keywords are used regardless of their differences in learning competency and knowledge level in that particular subject. This situation hinders the optimised use of Web search engines in finding relevant learning materials that match students' individual profiles. In this study, a Personalised Web search approach for E-learning is proposed. This proposed system augments the Web search engine. It provides recommendations of search results to students by using the group-based recommendation approach. The proposed approach is able to recommend results which match the students' learning competencies and behaviours. To evaluate the effectiveness and acceptance of the proposed system, an experiment was conducted among students. The results from the experiment suggest that the proposed approach created a notable improvement in terms of performance and satisfaction for the students.

Keywords: E-Learning, group-based recommendation, personalised Web Search, recommender system, student profiling.

PERSONALISED WEB SEARCH FOR E-LEARNING USING GROUP-BASED RECOMMENDATION APPROACH

ABSTRAK

Peningkatan pergantungan pelajar pada Web bagi tujuan pembelajaran adalah didorong oleh kewujudan dan pertumbuhan Internet. Enjin-enjin carian Web popular yang terdapat di pasaran bergantung pada penggunaan kata kunci yang tepat untuk mencari bahan pembelajaran yang relevan tanpa mengambil kira kecekapan pengguna mereka dari segi pembelajaran. Akibatnya, pelajar-pelajar akan menerima satu set hasil carian yang sama apabila kata kunci yang sama digunakan tanpa mengambil kira perbezaan dalam tahap kecekapan pembelajaran dan pengetahuan mereka dalam subjek tertentu. Keadaan ini menghalang penggunaan enjin carian Web secara optimum untuk mencari bahan pembelajaran yang relevan dan sepadan dengan profil individu pelajar. Dalam kajian ini, suatu pendekatan carian Web Peribadi bagi tujuan E-pembelajaran telah dicadangkan. Sistem yang dicadangkan ini menambah baik enjin carian Web. Sistem tersebut memberikan cadangan hasil carian kepada pelajar dengan menggunakan pendekatan cadangan berasaskan kumpulan. Pendekatan yang dicadangkan ini dapat membekalkan hasil carian yang sepadan dengan kecekapan dan tingkah-laku pembelajaran pelajar. Untuk menilai keberkesanan dan penerimaan sistem yang dicadangkan, satu eksperimen dijalankan dalam kalangan pelajar. Hasil daripada eksperimen menunjukkan bahawa pendekatan yang dicadangkan menunjukkan peningkatan yang ketara dari segi prestasi dan kepuasan dalam kalangan pelajar.

Keywords: E-Learning, group-based recommendation, personalised Web Search, recommender system, student profiling.

ACKNOWLEDGEMENT

Alhamdulillah, first of all, may all praises be to Allah, the most Beneficent and the most Merciful, the Exalted, the most Gracious, the Absolute Holy, the Victories, the Compeller, the Knower of all, who enabled me to undertake and complete my research work and to finally, write up the dissertation, leading to the fulfilment of a Masters of Computer Science degree.

Next, I would like to express my profound sense of gratitude to my supervisor, Associate Professor Dr Nor Aniza Abdullah for her inspiration, constant guidance, valuable suggestions, sympathetic advice and unparalleled encouragement given to me throughout the whole journey of this academic pursuit. She supported me morally and was often by my side as my mentor in every aspect of my academic journey as well as other issues in life.

I would also like to offer special thanks to my parents, my mother-in-law who always provided mental support and prayed for me throughout my academic pursuit and also to all the academic staff and friends from the University Malaya who supported me throughout my course of study.

Last but not the least, I offer my deep gratitude to my wife, Nowshin Zaman for staying with me at every moment of my academic pursuit since day one. Without her gentle encouragement, inspiration, support and love, it would not have been possible to complete this research and this dissertation.

Again, I am grateful to ALLAH swt, the Supreme Authority of the Universe for this entire research work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

LO	:	Learning Object
MIS	:	Management Information System
KP	:	Knowledge Point
TAM	:	Technology Acceptance Model
ANOVA	:	Analysis of variance

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Overview

Whether it is for achieving a particular job, acquiring knowledge or self-improvement, learning has always been a crucial part of human life since time immemorial. The massive amount of monetary resources invested in education by both the government and individuals clearly show the growing importance of learning in today's world, particularly for the improvement of the socioeconomic condition of society. With this in mind, human beings have always sought ways to improve their educational experiences. The ever-growing desire to obtain more effective and advanced learning methodologies, coupled with the increasingly competitive education industry, have motivated researchers as well as industry practitioners to develop more compelling learning environments whose aim is to enhance students' learning comprehension, thereby removing geographical and financial barriers. This phenomenon facilitates more flexibility for learning to take place from anywhere at anytime. The rapid advancement of the computer and multimedia technologies have inevitably provided educationists with the necessary resources to achieve the learning approach. (White, 2008).

In this modern era, the Internet offers learners or students immediate access to specific information anywhere at anytime, with minimal costs (Hennessy et al., 2010; Ingleby, 2012; White, 2008). Termed as "e-Learning", the Internet-based learning approach has transformed the way people learn, think and search for desired educational materials (Clark & Mayer, 2011). One only needs an Internet connection and a wireless device to get whatever information he wants. All that a user needs to do is to search for the topic through the Web search engine using a Web browser with the right combination of keywords. However, as it will be explained below, current available prominent commercial search engines such as Yahoo, Google, and Bing do not cater to the individual differences of learners in terms of learning styles, learning behaviours, outlook and learning capabilities (Curlango-Rosas, Ponce & Lopez-Morteo, 2011). This can adversely affect the learning process of the learners since the outcome generates the same learning materials to all students regardless of their learning capability, thereby deteriorating their learning performances (Felder & Brent, 2005; Premlatha, Dharani & Geetha, 2016). Based on this claim, there is a need for a Personalised Web search approach for e-Learning to be developed.

1.2 Research Problem

There are a number of ways in which the e-Learning mode practised today negatively affects the learning process of its users. First of all, the major commercial search engines available in the current market generate the same search results for all learners who typed in the same keywords. These search engines do not consider the learners' learning proficiency or level of experience in the particular topic (Curlango-Rosas, Ponce & Lopez-Morteo, 2011; Li, Luo & Mei, 2014; Premlatha et al., 2016; Wang & Wong, 2013). Consequently, this makes learning extremely difficult for many learners as they struggle to find the desired relevant resources that could match their learning comprehension ability. For instance, a video tutorial created for advanced learners may not be a good fit for learners who are novice. This problem is further aggravated by the fact that many learners, particularly novice learners, struggle a lot to find the right combinations of keywords to express their need in their search query even though relevance of the search results strongly depends on the right combination of keywords (Curlango-Rosas et al., 2011; Kumar & Ashraf, 2015; S, K & G.S, 2014; Yathongchai, Angskun, Yathongchai & Angskun, 2013). Additionally, with the huge number of search results returned by the search engines, learners often find it challenging to select the most relevant ones which will meet their needs (Capra, Arguello, Crescenzi & Vardell, 2015; Capra, Marchionini, Oh, Stutzman & Zhang, 2007; Curlango-Rosas et al., 2011; Premlatha et al., 2016). In fact, research (Hassan & Mihalcea, 2011) has illustrated that while looking for any educational materials, a dominant search engine usually generates only four results out of the top 50 results that can serve the educative purpose. Furthermore, the major search engines actually offer services as business entities with the aim of making profits and they often place irrelevant advertisements in the form of CTR (Click-through rates), CPM (Cost per thousand viewers), CPA (Cost per action), CPC (Cost per click), and others on top of the more relevant links in their search results. This strategy creates more confusion among the learners as they struggle to select the appropriate links (Curlango-Rosas et al., 2011; Iverson, 2011; Zweihorn, 2006).

Even though search engines are conventionally evaluated based on the relevance of the returned search results in response to individual queries (Arbor, Arbor & Jones, 2010), such evaluations still do not give a complete picture as to how well the search engines meet the users' requirements. This is because the same search query may carry different meanings regardless of the user's intention. For instance, the keyword, "Apple", carries various meanings. It could refer to Apple Inc. or the real fruit. A previous study (Tamine & Pierre, 2016) showed that the relevance of search results does not necessarily imply users' satisfaction. A better evaluation would be based on how well the search results can fulfil the users' personal needs and personal assessment on the usefulness of the returned search results (Tamine & Pierre, 2016). Additionally, similar evaluations can be performed in a group learning environment where the assessments gained from a group of similar prospective users are taken into account (Liu, 2006).

The abovementioned discussion shows that there is still room for improvement in the search engine industry, particularly in addressing the educational needs of its users. One way to improve the search engine is to align the users' profile information (which will reveal users' learning capability) with the submitted queries. In this manner, the search engine would become personalised by satisfying the users' requirements more effectively. In fact, personalised systems which take into account user preferences, interests and browsing behaviours, already existed on the Internet in various forms of the Web applications, with an aim to provide personalised services to the users (Chen, Lee & Chen, 2005). For example, there are specialised search applications available in the market that can assist users in their travel, shopping, entertainment or personal needs such as Agoda.com recommends tourism spots and hotels, and YouTube recommends video. Nevertheless, research (Qiu & Cho, 2006; Amershi & Morris, 2008; Anuradha, 2012; Morris, Teevan & Bush, 2008; Morris, 2008; Tamine & Pierre, 2016; Teevan, Morris & Bush, 2009) showed that personalisation algorithms perform best with huge amounts of data about an individual, some of which can be acquired by using data from other people. Besides, the outcome of the personalisation system can also be refined by using the behavioural information of a specific group of people, especially for explicit groups and group related queries (Amer-Yahia, Roy & Chawlat, 2009; Amershi & Morris, 2008; Ahmed, Nabli & Gargouri, 2012; Quijano-Sánchezz, Recio-García & Díaz-Agudo, 2011). Another important information that can be used to improve users' searching experience is the way other users approach similar tasks, previously (Capra et al., 2015). The main idea is to combine an individual's personal data with the information gathered from other related groups of people in order to enhance the performance of the personalised system that will lead to better searching experiences.

The importance of the personalised system in e-Learning was highlighted by a study (Capra et al., 2015) which showed that when users are exploring for educational materials on unfamiliar topics their feeling towards the urge for guidance while searching gets stronger. Although the personalised system is available in the market for various purposes such as entertainment and leisure, personalised recommendations for suitable e-Learning materials based on learners' competency is still lacking when using popular search engines, despite the urgent need.

Based on this issue, there is a scope to improve learners' e-Learning comprehension via better personalised Web search experiences by using the group related information from similar groups which can then leverage students into searching for learning materials more easily.

1.3 Research Questions

From the aims and objectives defined, the research questions formulated for this research encompasses:

1. What are the problems encountered by e-Learners while using the existing Web search engines for e-learning?
2. What are the features offered by the specialised Web search engines for e-Learners?
3. How to provide personalised recommendations of Web search results to heterogeneous e-Learners who have different learning capabilities and needs?
4. How to dynamically model the needs of each individual e-Learner?
5. How to determine e-Learners' satisfaction with the proposed system?

1.4 Research Aim and Objectives

Based on the discussion provided above, the aim of this research is to propose a Personalised Web search approach for students of different learning capabilities in e-Learning by using the group-based recommendation approach. Hence, the objectives of this study are:

1) To investigate the strengths and weaknesses of the existing Web search in order to accommodate the needs of individual students for searching e-Learning materials.
2) To design a personalised Web search approach for e-Learning by using dynamic learner's profiling and a group-based recommendation technique.
3) To evaluate the feasibility of the proposed Web search approach for e-Learning.

1.5 Scope of the Research

This research work focusses on the approach of Personalised Web search for e-Learning by using the group-based recommendation approach. Due to the time restriction in accomplishing the research objectives, the scope of this research is thus limited to the following:

i. The proposed Personalised Web search engine only interfaces with the Google search engine.
ii. The target users are undergraduate university students; therefore, teachers and the general public are not taken into account.
iii. Students present and previous academic records, Web browsing histories and session data are considered as a dataset to evaluate the proposed method.
iv. The feasibility of the approach is evaluated based on the performance and the satisfaction of the students towards the proposed approach.

1.6 Research Significance

The enormity and diversity of the Web contents have necessitated revisiting the concept of ‘one search result fits for all’ proposition of the contemporary Web search engines. Currently, there are reliable and personalised search engines in the market to facilitate users cater to their daily necessities such as shopping, entertainment, personal requirements and travelling. However, we hold that one community namely the 'students' particularly requires urgent attention regarding recommendations for personalised e-Learning materials from their Web searches. This study aims to address and fill up this gap. This study thus attempts to understand e-Learners’ personal learning needs, their learning behaviours and their group relationships while using Web search engines for their learning purposes. Consequently, this understanding is translated into a personalised Web search tool that can enable e-Learners of different learning competencies to receive recommended Web search results that match their learning profiles.

1.7 Organisation of the Thesis

There are seven chapters in this dissertation. Chapter 1 briefly describes the background on the uses of the Web search engines to retrieve the desired contents. Using a Web search engine as an e-Learning tool to seek learning materials is also discussed. The motivation of the work, problem statement, research aim and objective, and scope of the work are explained. The contribution of this study and the highlights gained from the proposed approach are mentioned at the end of the chapter.

Chapter 2 begins with the discussion of the Web search engines. It briefly explains personalised Web search including features and components required, and techniques used for delivering personalised search results. In relations to that, the chapter provides a discussion of e-Learning which includes Web search in e-Learning and analysis of various specialised Web search engines used in e-Learning along with their strength and weakness. It also studies several factors that need to be considered while delivering personalised e-Learning materials. The overview of the recommender system, and its general architecture and techniques used in the recommendation generation process are also discussed for delivering personalised Web search recommendations for e-Learners. A chapter summary is presented at the end of the chapter.

Chapter 3 illustrates the methodology used in this research work. The details of the data collection techniques, materials and instruments used, design of the proposed framework, implementation of strategies, and evaluation methods are discussed.

Chapter 4 begins with a discussion on the overview of the proposed framework of a personalised Web search for e-Learning by using group-based recommendations. Each of the components required to form the framework and the mechanisms involved in each of the components is explained in detail. A chapter summary is presented at the end of the chapter.

Chapter 5 discusses the design and implementation of the proposed framework. First, the architectural design of the proposed system is explained briefly. Next, the designs of the database and user interface are presented. The implementation mechanism of the proposed system is analysed at the end of the chapter.

Chapter 6 explains the experimental design, the experimental procedure, the data collection, and the evaluation of the proposed system. It explicates the evaluation results, the analysis and findings of the proposed approach in detail. It also briefly discusses the outcome of the approach in terms of performance and satisfaction of the students at the end of the chapter.

Chapter 7 concludes the research work and its achievements. It also presents some limitations and recommendations for future work.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter begins with the introduction of Web search engines. A detailed analysis of the personalised Web search is performed that includes its system components, requirements and techniques. The discussion on e-Learning and the use of search engines in e-Learning along with the strength and weakness of the various specialised Web search engines for e-Learners are also elaborated. Furthermore, the recommender system and its general architecture are analysed. Different techniques used in delivering the recommendations are also reported. The chapter concludes with a discussion on the personalised recommendation for Web search in e-Learning. A chapter summary is presented at the end of the chapter.

2.2 Web Search Engines

Millions of people all over the world are using Web Search engines every single moment. The term, search engine, usually refers to the well known commercial Web search engines such as Google, Bing, and Yahoo. It also refers to a broad array of search systems, for instance, email, social networks, and commercial service providers that are part of the mainstream Web-based applications (Marin, Gil-Costa, Bonacic & Inostrosa, 2017). Three principal elements usually make the Web search engines run: the crawler retrieves the documents from the Web, the indexer indexes the documents gathered by the crawler, and the searcher submits user enquiries to the search engine database to achieve expected search results. The below Figure 2.1 illustrates the relationship within these three elements. In general, search engines are used to answer information needs. They are large databases of software packages and Web pages which are meant to index and retrieve the pages and then enable the users to find their required information.

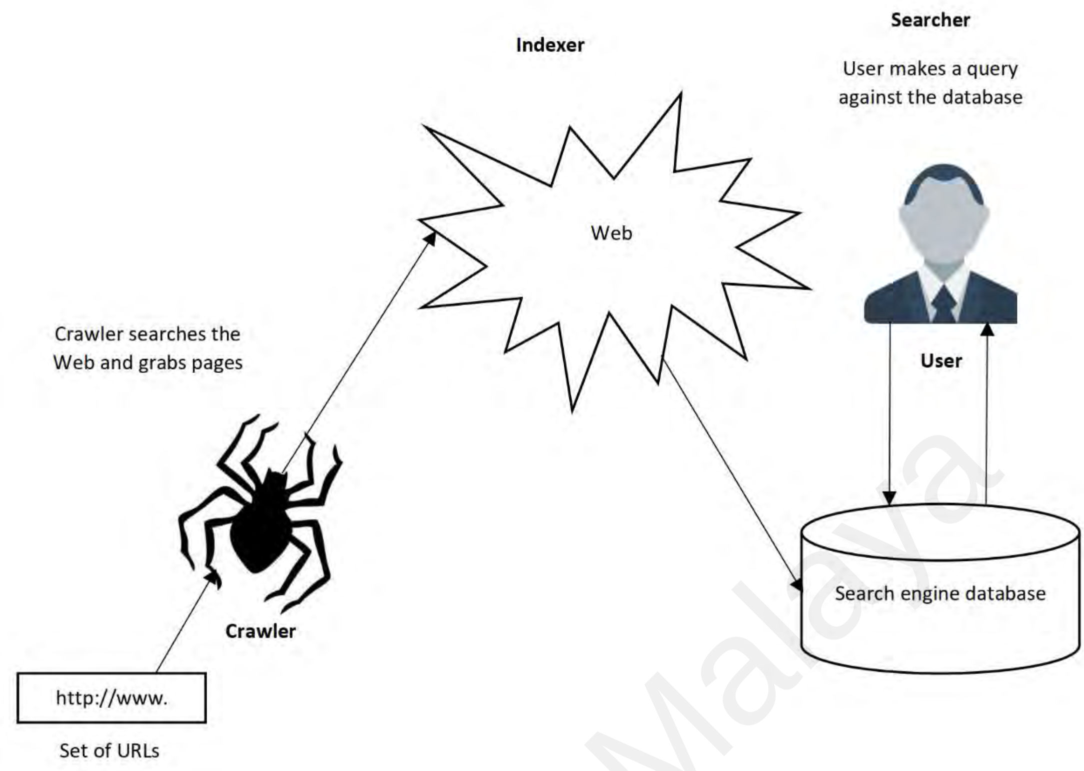

Figure 2.1: Overview of Web search engine

As Figure 2.1 demonstrates, users express their information needs as queries using keywords and the Web search engine then returns a list of results based on the submitted keywords. It takes into consideration the users' intention and information needs; typically, ten search results are generated per page (Curlango-Rosas et al., 2011). In most of the existing Web search engines, search results are retrieved by evaluating the relative importance of the links. Ranking algorithms calculate the rank of a webpage by considering keywords, high-quality contents and backlinks from external websites that flow naturally. The search Web engines use this information and determine the rank of the Websites.

2.3 Personalised Web Search

A user query generated by a user creates a vast number of Web pages returned by the Web search engines, many of which may be irrelevant and ambiguous to the searcher

(Keenoy & Levene, 2005; Li et al., 2014; Priyanka & Vinod, 2014). Different users have different background and information needs, and the Web search engines need to address the variance in the informational goals of users using the search engines in a personalised way (Teevan, Dumais & Horvitz, 2005). Studies have shown that even for the same query, people differ significantly in terms of the search results they deem appropriate, particularly, when their information needs are different from each other and when they display the underlying intention for the query in similar ways (Teevan, Dumais & Horvitz, 2005). For example, students searching for "LinkedList in java" will receive a huge number of search results from the Web search engines. The order of the returned results from the search engine will be the same for all students regardless of their learning proficiency. Therefore, the search results need to be customised differently for students with different levels of learning competencies such as novice learners or advanced learners. A personal Web search engine should consider its users' intentions, behaviours, preferences and background before delivering the personalised search results to the users.

2.3.1 System Components of the Personalized Web Search

A Personalised Web search adopts a personalisation architecture. It attempts to provide results that suit the interest of the user as an individual or as a member of a group. Priyanka and Vinod (2014) presented a general architecture of the Personalised Web search, as presented in Figure 2.2. The Personalized Web application module works as a middleware between the Web search engine and the Web user. It obtains user requirements (preferences and interest) from the user and stores these in the user profile module.

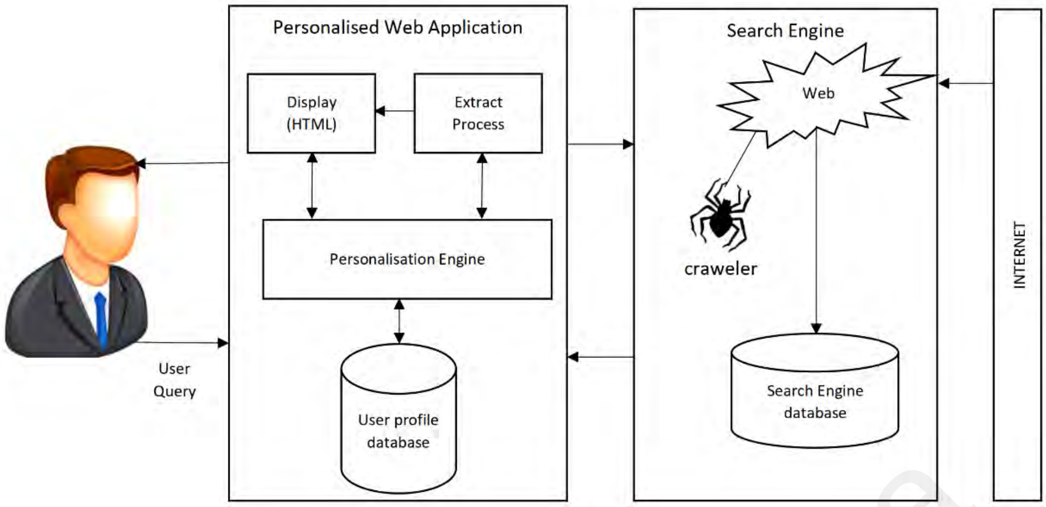

Figure 2.2: General architecture of the Personalised Web search (Priyanka & Vinod, 2014)

Subsequently, it sends the user's query to the search engine. The returned search results generated by the search engine is received by the Personalised Web application module for further processing so as to personalise it for the user. Before returning the customised results to the user, the module uses different personalisation techniques to process the extracted search results inside the personalisation engine.

2.3.2 Requirements of the Personalized Web Search System

Keenoy and Levene (2005) have identified an ideal personalised search system. They assert that such a system should exhibit some distinguishing features such as:

- **User data collection method**: The data should either be collected explicitly from the users or implicitly deduced from their normal interactions with the system.
- **Profile storage:** User's data can be stored client-side on the user's machine or at a server.
- **Adaptivity**: Over time, the system adapts automatically as per the user's requirement. Hence, a personalisation system needs to be capable of identifying the changes in the user's context such as interests and preferences

and it should reflect the changes in the system in order to deliver the personalised contents dynamically.

- **Types of data**: Several types of data are collected following location, search history, preferences, Web history, community connectedness and others.
- **Algorithm**: Different types of algorithm(s) are used depending on the context of the personalisation (such as collaborative filtering, hybrid filtering) so as to be able to re-rank the search results and to present them in a personalised way.
- **Interface**: Personalised results could be presented across different platforms and devices such as mobiles, desktops, tablets and laptops via Web browsers and applications.

2.3.3 Techniques Used in the Personalised Web Search

In order to deliver the Personalised Web search results that consider the users' requirements and intentions, previous researchers have introduced several techniques as reported in the literature. Those techniques can be clustered into relevant groups. Some of the relevant techniques are discussed below.

2.3.3.1 Query-based Personalisation

i. **Identifying semantic similarity between query words**: Personalisation of Web search results can be achieved by computing and analysing the semantic relationships between user's search queries (Makvana, Jay, Shah, & Thakkar, 2016). To accomplish this, personalised ontology from the associated user's previous queries and session logs is created that helps to identify the semantic similarity relation within the query words. The underlying relation between query context/topic investigated helps to improve evaluation measures used for calculating semantic relatedness between words which results in improvement of search result ranking.

ii. **Query-aware attention**: The order of the issued search queries is significant to discover the real user interest along with other information such as user's query, past clicks, topical interests, click entropy and the sequence of previous queries and sessions in order to tailor the original Web search results ranking (Ge, Dou, Jiang, Nie, & Wen, 2018). Compared to the older search sessions the newest sessions may deliver more reliable personal signals. Moreover, the personalisation of the current query is also influenced by previous search histories and user behaviours. Therefore, to exploit such sequential information, Ge and Dou (2018) proposed a query-aware attention model using the hierarchical recurrent neural network which is capable of generating dynamic user profile automatically from historical data that is aware of the input query. The technique is able to improve the search result ranking by scrutinising all prior sessions and highlighting the more relevant sessions in a dynamic way with a view to presenting the information need after fine training.

2.3.3.2 User Profiling-based Personalisation

i. **Concept-based user profiles**: Here, an extended set of conceptual preferences is acquired for a user which is based on the concepts obtained from the clickthrough data results from the Web search (Leung, Lee, Ng, & Fung, 2012). Then, a concept-based user profile (CUP) which represents the user profile is generated as a concept ontology tree. The ontology enables the technique to gather rich user concept preferences, in addition to those straight away derived from the user clickthroughs. Finally, the CUP is keyed into a support vector machine (SVM). The aim is to learn a concept preference vector to adopt a personalised ranking function that can re-rank the search results.

ii. **Fuzzy based user profile classification**: The fuzzification functions play a major role in handling uncertainty data in vague environments by classifying web users in a personalised search setup. However, this is complicated due to the very nature of dynamism that exists in user browsing history. This fluctuating nature of user behaviour and user interest can be interpreted within a fuzzy setting deftly. Sendhilkumar and Selvakumar (2014) proposed a fuzzy based user classification model to leverage the Personalised Web search environment. The web browsing data are fuzzified and fuzzy rules are created using decision trees. Consequently, through fuzzy rules, the search pages are labelled for the purpose of identifying the level of the Web users interest towards searching in order to present a personalised search result.

iii. **Temporal-based user profile modelling**: Usually users submit queries to Web search engines that contain their information needs. In return, a list of Web links is returned as search results to them that they can be interested in. If the returned information is unable to satisfy the needs of a user, he/she restructures the previous query in a different way (Kacem, Boughanem, & Faiz, 2017). Therefore, it is essential to understand users' interest and preferences from the submitted queries in a session (Kacem et al., 2017). A current query in a session can be modelled via a temporal-based user profile and expressed prior interactions (such as clicked results, reformulated queries and submitted queries). The user profile impacts on the accuracy of the search result in the context of session search by combining both freshness and frequency of the user's actions in a time-sensitive manner.

iv. **Personalisation of search results with editable profiles**: Personalised search results is essential to improve the search performance. The existing Web search engines do not offer the authority to customise the returned search results (Zemede

& Gao, 2017). The lack of flexibility and user's command and may frequently cause inconvenience and non-productive experience for users. To overcome the shortcomings, a transparent search personalisation technique titled PEEPLER is proposed that facilitates users with full control and manipulation of search results (Zemede & Gao, 2017). In PEEPLER, a user can own more than one profile, and each of these can be altered arbitrarily. Profile terms are either entered manually or generated automatically and supplemented by adding the ones that are semantically related. Negatives terms are also considered that allows filtering unwanted search results. The chosen profile enables PEEPLER to re-rank the search results based on their consistency to the profile.

2.3.3.3 Group-based Personalisation

i. **Community-aware**: Social media reveals significant information about the interest of a user and his/her community. Exploiting this information may leverage the ranking of the Web search results (Sarker et al., 2015). To achieve this, Shafiq, Alhajj and Rokne (2010) proposed a community relation based approach to find out activities of a user's social network and what information the user obtains from the social networks. Based on the information regarding the community connectedness the results of a Web search engine are prioritised.

ii. **Exploiting similar users' behaviour**: Users get benefited from the search experience of other users with alike interests and other features of searching when they search for the repetitive queries (Y.-J. Li, Li, & Lin, 2017). Therefore, it is possible to rearrange the search results in a way that the most relative results from similar users will be placed ahead of the result page. To attain this, first users are clustered based on their demographic information and search features using DBSCAN (Density-Based Spatial Clustering of

Applications with Noise) clustering method. Secondly, filtering the irrelative results by analysing the click log of previous users. Consequently, top relevant results will be placed ahead of the result page as recommendation results.

iii. **Discovering and using groups to improve personalisation**: Exploiting community information by discovering the users' relationships among the members in a group is one of the ways to surpass the ranking of the personalised Web search results (Sathiyabama & Vivekanandan, 2011). This can be achieved by analysing community connectedness. It is an uncomplicated way to acquire more accurate information in accordance with their interests and preferences of the group members. Therefore, it assists in providing more relevant information about a particular user by identifying his/her interest and then personalising the search results accordingly. Groupization is a personalisation technique that combines personal and group related information which is then used to improve the Web ranking for different group/query combinations (Teevan et al., 2009). It considers data related to all group members despite data connections with a single user (Ahmed et al., 2012). The technique analyses the similarity of the query choices, personal contents, and relevance judgments for different categories of implicitly and explicitly-defined trait-based and task-based groups so as to enhance the effectiveness of personalisation of the Web search results.

2.3.3.4 Context-based Personalisation

i. **Employing situational context to improve personalisation**: Situational context has a dynamic role to improve the ranking quality of the personalised search (Zamani, Bendersky, Wang, & Zhang, 2017). Therefore, it is required to consider the situational context as a property of the current search request that is independent of both query content and user search histories such as location and time. To accomplish this, one way could be designing context-aware ranking models based on neural networks by using click data that is collected from the users' Web search histories in order to improve rankings in different scenarios (Zamani et al., 2017). The identified relationship between the user behaviour and situational context from the search logs may leverage the personalised search.

Various types of personalisation techniques have been discussed in section 2.3.3. In the mentioned studies, researchers have focused on three primary aspects of personalisation to satisfy user's personal needs. First is in the query-based technique, the semantic relationship between issued search query words is attempted to discover along with their order to improve the search result ranking. Secondly, several researchers pointed out the significance of user profiling in the personalisation of Web search results while considering users' interest and preferences, clickthrough data, browsing history, reformulated queries and authoritativeness to customise the returned search results. In user profiling, user's community such as groups and social media may also reveal valuable insights on identifying user's relationship with other similar minded users that can provide more accurate information regarding user's interest profile to improve the effectiveness of a personalised search engine. Lastly, in Web search session, the contextual information such as location and time also plays a key role in identifying user's situation and intention which may surpass the search result ranking. Based on the discovered knowledge on various personalisation techniques it is understood that a Web search engine primarily needs to focus on dynamic user profiling method while information regarding community connectedness and situational context are taken into consideration. This will supplement the retrieved personal information for a better understanding of the user. Additionally, the investigated semantic relation between submitted query words may also help the search engine to understand exactly what users' mean and give back exactly what they want. This knowledge will bring significant insight when designing any personalised Web search recommender system.

2.4 E-Learning

In general, Electronic learning or e-Learning refers to using technology to acquire education or training (Rizk, Gheith & Nasr, 2017). It includes both the technical support to learning and the other media and resources, for instance, satellite broadcasting of lectures, video, interactive television, intranet, wireless and mobile devices so on and so forth (Klašnja-Milićević, Vesin, Ivanović, Budimac, & Jain, 2017). However, e-Learning is converted into internet learning as it primarily uses internet technologies for the creation, adoption, transfer and facilitation of the learning process. Most of the cases, it refers to a program or course or degree delivered entirely online.

Typically, e-Learning courses are particularly provided via internet instead of the classroom where the instructor is teaching. It can be defined as an interactive online teaching process which allows learners to communicate with teachers or instructors and other students. Sometimes it is delivered in real time where students can raise their hands and interact in real time, and sometimes it is delivered through pre-recorded lectures.

E-learning systems are presently becoming an integral part of the educational and business organisations due to the introduction of computers and the internet in the late 20th century (Executive Summary, 2013). It creates a virtual learning environment that enables people to gain access to the richness of online information and new e-Learning opportunities (Klašnja-Milićević et al., 2017). Moreover, in recent times a lot of learning management systems (LMS) have been developed which is a concept emerged directly from e-Learning. LMSs are primarily focused on online learning. However, they support a range of uses, acts as a platform for online content such as courses including both asynchronous based and synchronous based. For example, academia open source MOODLE system has been in use all over the world (Marcato & Scala, 2012). It facilitates students and teachers with full functionalities such as exchange learning materials, performing tests, communicate with each other in various ways, track and trace the progress, and so on. According to the U.S. higher education market report of fall 2018, the top three most accepted LMSs (by the number of institutions) are Blackboard (31%), Canvas (30%), and Moodle (18%) (6th Annual LMS Data Update | edutechnica, 2018). The same three systems led in terms of the number of students enrolled; however, Canvas slightly surpassed Blackboard.

2.4.1 Web Search Engines for e-Learning

In current trends, the e-Learning Systems enable learners of all ages, preferences, and competencies who are looking for knowledge or information "anywhere at anytime" (Suguna, Sundaravadivelu & Gomathi, 2016). It provides instant access to learners to fulfil their respective needs. E-Learning systems can be employed either as Web-based systems for online education or an additional method for on-campus study, serving as a supportive learning tool for students such as Web search engines. Study has found that among all e-Learning tools, around 80% of students prefer Web search engines on the internet to meet their educational needs (Dogruer, Eyyam, & Menevis, 2011). This indicates the significance of Web search engines as an e-Learning tool among students, other than their classroom learning environment. With the help of Web search engines, teachers can easily find relevant resources and materials while students can use relevant search engines to seek information and to access multimedia materials to complete their assignments. Relevant search engines also enable learners to gather knowledge from different domains (Curlango-Rosas et al., 2011). Nevertheless, learning styles, behaviours, outlooks and potentials vary from one student to another and without doubt, the learning process is influenced by these differences. Viewed from that perspective, research has revealed that providing the same learning contents to all types of learners with different learning competencies may not provide the same learning experiences. It may even decrease their learning performances (Premlatha et al., 2016).

To access the Web-based learning system, learners type the keywords into any major commercial search engines such as Google, Yahoo, and Bing. This search tends to generate the same set of search results for the same set of submitted keywords regardless of the learners' learning proficiencies (Curlango-Rosas et al., 2011; Ge et al., 2018; Y.-J. Li et al., 2017; Makvana et al., 2016). Moreover, it was discovered that only four results out of the top 50 returned results that are received using the major Web search engines turn out to be profoundly educative (Curlango-Rosas et al., 2011). Nevertheless, these search engines only considered the right combination of keywords submitted by learners as a measurement of relevancy in order to return suitable search results. This phenomenon affects the novice learners who will especially find the process difficult since they may not know the right keywords to type when they are new to the topic (Curlango-Rosas et al., 2011; Kumar & Ashraf, 2015; Yathongchai et al., 2013). Additionally, students also find it demanding to choose the most appropriate links from the huge number of links which the Web search engines return (Capra et al., 2015; Curlango-Rosas et al., 2011; Premlatha et al., 2016). Major search engines also tend to place commercial sponsored adverts links in the form of CPM (Cost per thousand viewers), CPC (Cost per click), CPA (Cost per action), or CTR (Click-through rates) on top of the relevant links and such diversions can distract learners who are making attempts to choose the suitable links from the returned search results (Curlango-Rosas et al., 2011; Sherman, 2001; Zenetti, Bijmolt, Leeflang & Klapper, 2014). In relation to this, search engines are regularly evaluated in terms of the relevancy of the Web pages to specific queries (Hassan et al., 2010) however, in reality, barely a small portion of the students' overall requirement for information is represented by the search request. Since learners may have diverse backgrounds and different expectations for a given query, it is only appropriate when these Web search engines personalise their results by considering the students' overall interests and preferences (Makvana, Shah, & Shah, 2014). Therefore, in e-Learning, it is necessary for the Web search engines to understand the students' learning needs and deliver the learning materials that are adequate to their profiles. The next section discusses different types of search engines reported in the literature (Capra et al., 2015; Curlango-Rosas et al., 2011; iSEEK - Education, 2007; Microsoft, 2013) which have been specially built for e-Learning purposes.

2.4.2 Specialised Web Search Engines for e-Learning

Traditional Web search engines have diverse specialised search engines for various domains. In general, specialised search engines are those that can deliver materials coming from any particular domain such as Google Scholar which delivers academic papers (Cecchino, 2010) and Agoda which delivers tourism-related materials (agoda.com) to be specific. In other words, Web search engines that possess some additional features or assistive components can influence traditional search engine users to find the required materials more easily (Curlango-Rosas et al., 2011). Likewise, specialised search engines for e-Learning that possess special features or assistive components can enhance e-learning by providing e-Learners with more support to search for educational materials with limited effort. The following subsections discuss several specialised search engines for e-Learning as reported in the literature (Capra et al., 2015; Curlango-Rosas et al., 2011; iSEEK - Education, 2007; Microsoft, 2013).

2.4.2.1 Learning Object Search Tool Enhancer (LOBSTER)

The Learning Object Search Tool Enhancer (LOBSTER) is a specialised search assistant tool. It works alongside the Google search engine. It was developed with an aim to help teachers to search for learning objects (LO) more easily in their

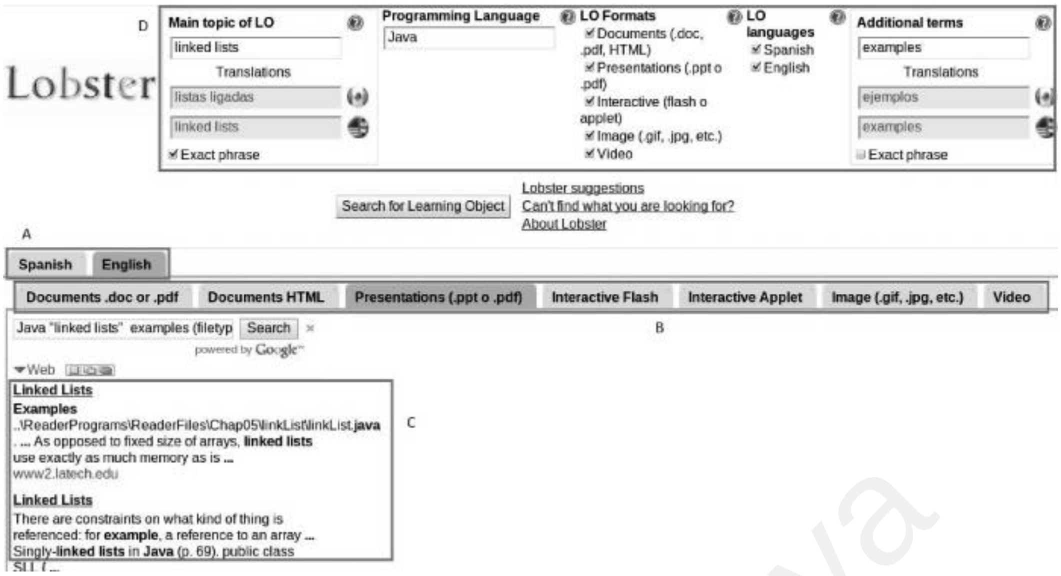

Figure 2.3: Search interface of LOBSTER (Curlango-Rosas et al., 2011)

Web search (Curlango-Rosas et al., 2011). The key contribution of the tool is providing a set of assistive components that support teachers to look for LOs by using the Google search engine in their entire search process. Figure 2.3 shows the user interface of the LOBSTER (Curlango-Rosas et al., 2011).

The search assistant tool comprises a set of features such as bilingual search and term suggestions which help to improve teachers' search experience. LOBSTER also offers additional support to users by offering bilingual topic-specific term suggestions, clustering of search results based on language and LO types, advanced searches, as well as suggestions based on appropriate query terms. In fact, Curlango-Rosas (2011) has shown that these assistive components and additional features helped users to search for the learning objects more successfully than that of the direct search of the Google search engine. LOBSTER significantly increases the number of times that teachers find their desired LOs. However, there are still some limitations in this tool. The system architecture does not accommodate the user profiling and content re-ranking modules that can identify the individual learner's need and interest to re-rank the search results based on individual profile.

As a consequence, the system returns the same results for the same query terms without taking individual differences into account such as user's learning competency.

2.4.2.2 Search Guide (SG) by Displaying Search Trails

Research (Capra et al., 2015) has shown that displaying search trials from previous users who were looking for similar Web contents may assist new users in their Web search. Capra and Arguello (2015) thus developed a novel tool called the search guide (SG). It was based on the Microsoft Bing search engine which shows the search trails (such as queries issued, pages bookmarked, results clicked etc.) from three prior users who have accomplished the same task. These additional features provided by the Bing search engine enabled the users to perform more complex search tasks with ease rather than just using the traditional Web search engine, as shown in Figure 2.4.

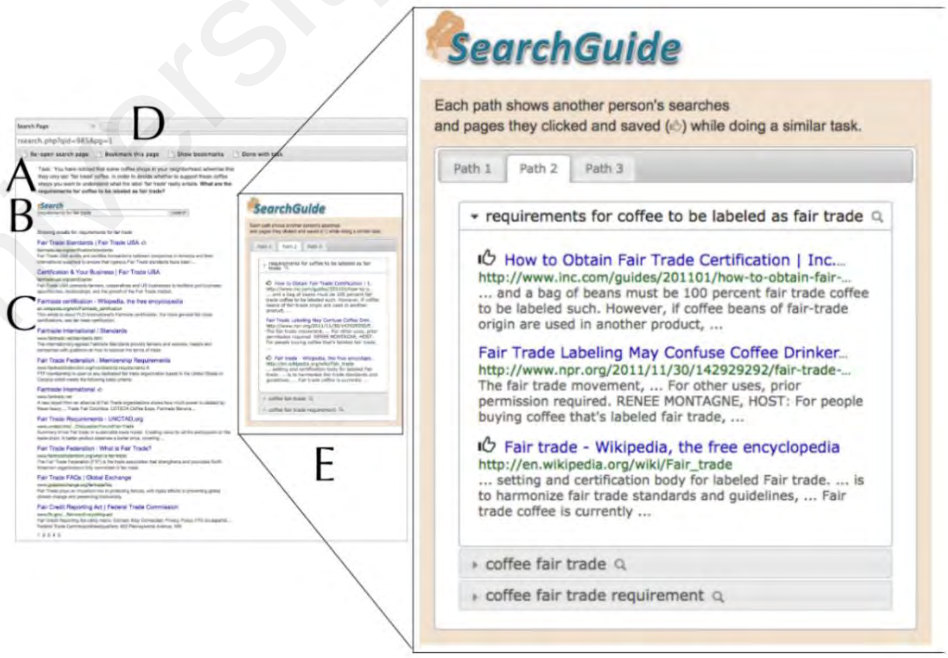

Figure 2.4: Search interface of search guide (Capra et al., 2015)

The study (Capra & Arguello, 2015) revealed that when search trial components were inserted into traditional search engines as an additional feature, the Web search users searching experiences improved. Thus, the search guide enabled the users to find their desired materials with limited effort, particularly when the task is complex. However, the search guide is unable to personalise the Web search experience in accordance with the users' personal learning profile as it does not contain learners profile information. In addition, it presents the same search results as traditional search engines do due to the absence of ranking mechanism.

2.4.2.3 iSEEK Education

The iSEEK Education is an academic search engine (Figure 2.5) which is geared towards helping users find reliable academic materials such as term papers, research projects, and anything that requires reliable citations (iSEEK - Education, 2007). It is intended for educators, teachers, and pupils to assist them in finding relevant and high-quality results on the web. This search engine gathers numerous authoritative resources from government bodies, universities, and prominent non-commercial providers. The results are safeguarded to protect children and it is viruses free. Moreover, it confirms the generated results to be trustworthy and reputable. The authoritative results are accumulated from copious trusted resources, that are of top quality and reliably reviewed by leading educators. As this search engine allows users to view the information which iSEEK has gathered so it enables more targeted results to be shared thereby, eliminating a huge amount of irrelevant information. However, like LOBSTER, the iSEEK system does not address dissimilarity among users (such as learning styles and preferences) due to lack of personalisation features. It also presents the same materials to different users. A notable limitation of the iSEEK's search process is that it is unable to

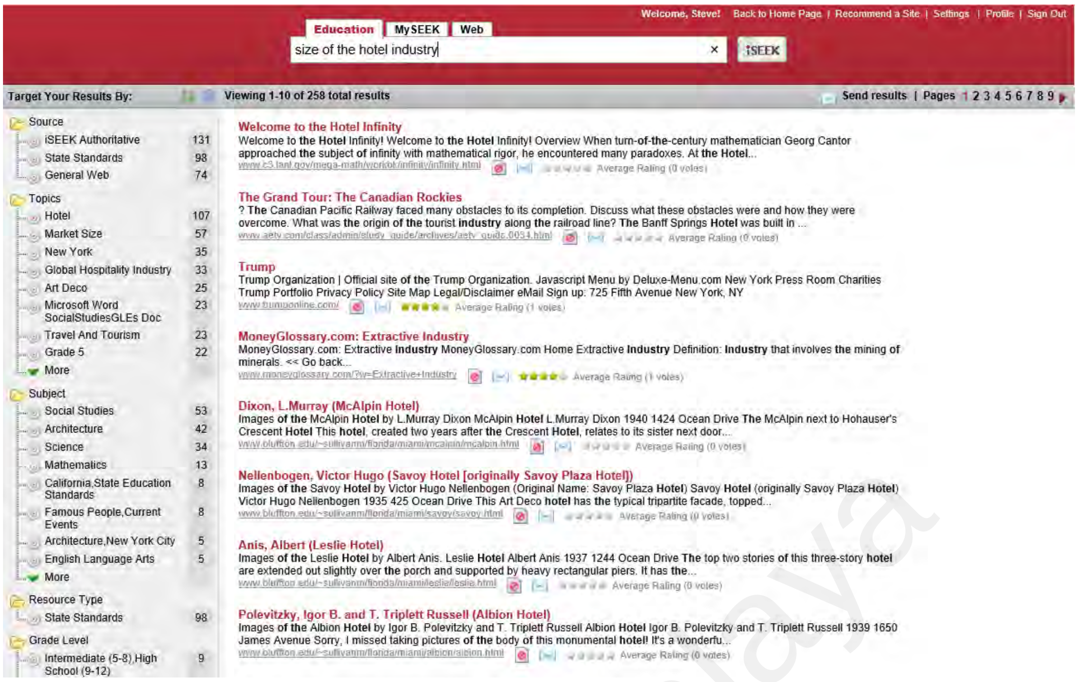

Figure 2.5: Search interface of iSEEK Education (iSEEK - Education, 2007)

search the entire Web like traditional search engines (Google, Bing) do.

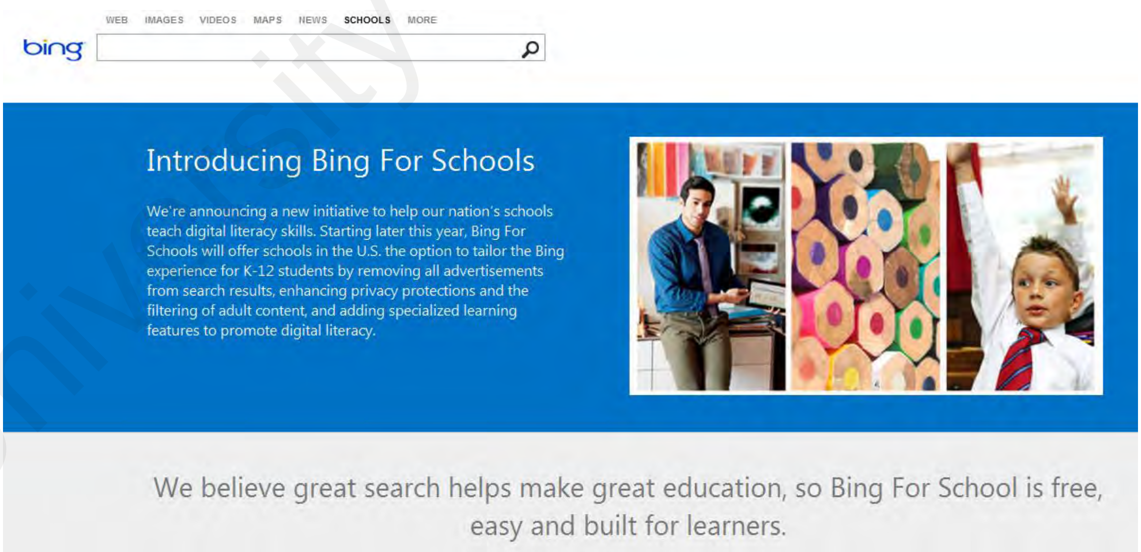

Figure 2.6: Search interface of Bing-for-Schools (Microsoft, 2013)

2.4.2.4 Microsoft "Bing-for-Schools" program

In 2013, an ad-free "Bing for School" program was launched by Microsoft with a view to providing K-12 students with an ad-free search experience. Figure 2.6 shows the interface of the Bing-for-Schools search engine. The program is equipped with additional content filters, privacy protections and other specialised learning features that could enhance digital literacy (Microsoft, 2013). Most importantly, the program aims to offer a distinct search engine that is designed specifically for educational purposes. Today, this program is called "Bing in the Classroom" – out of Beta. Opened to all K-12 schools in the US, Microsoft claimed that the program was able to attract schools in the five biggest districts in the US with 4.5 million students as users and serving around 35 million ad-free queries. Even though this means that only about eight queries were done by each student per day – which is not a very striking number since a majority of the people perhaps do more than eight Google searches per day, Microsoft stated that the program is now evolving by more than a million enquiries per day. To the best of our knowledge, Microsoft "Bing-for-Schools" program does not consider the personalisation of search results according to the learner profile. Table 2.1 summarises the strengths and weaknesses of the personalised search engines discussed above.

2.4.3 Personalisation in e-Learning

In e-Learning, 'personalisation' comes up with a broad range of new meanings. One of the best interpretation could be that "Personalized learning is the tailoring of pedagogy, curriculum and learning environments to meet the needs and learning styles of individual learners" (Baguley et al., 2014).

Table 2.1: Strengths and weakness of specialised search engines

Name	Strength	Weakness
LOBSTER	✓ Bilingual search and term suggestion. ✓ Clustering of search results based and language and LO type. ✓ Bilingual topic-specific term suggestion. ✓ Advanced searches as well as suggestions on appropriate query terms. ✓ Search the entire web/resources.	X Dynamic profiling. X Personalized recommendation. X Personalized search result.
SG	✓ Exhibits trails of searches (bookmarked pages, queries issued, results clicked) from three past users who performed the task ✓ Search the entire web/resources	X Dynamic profiling. X Personalized recommendation. X Personalized search result.
iSEEK Education	✓ Authoritative results are collected from numerous trusted resources reviewed by leading educators for greater quality and trustworthiness ✓ Eliminate lots of irrelevant information	X Dynamic profiling. X Personalized recommendation. X To search entire Web. X Personalized search result.
Bing-for-Schools	✓ An ad-free search experience for K-12 students ✓ Equipped with additional content filters, privacy protections ✓ Advance search features	X Dynamic profiling. X Personalized recommendation. X To search entire Web X Personalized search result.

Personalisation shifts the teacher-centric approach of teaching to a learner-centric, competency oriented one. In traditional approach one fit to all learning style where e-Learning recognises learners as a heterogeneous mix of individuals,whereas, personalised e-Learning offers customisation of a plethora of materials in the online education process such as learning content, learning environment, and the interaction. Apart from the personalised settings, there are other features of the learning environment that can be personalised, for example, types of the deliverable content and the style to be delivered, student's acceptance towards the approach etc.

In recent times, educators need to re-examine e-Learning courses where a plethora of essential factors such as background, age, culture, educational level and demographics information plays a significant role to determine it. Various significant factors need to be considered when deciding to personalise an e-Learning environment (Klašnja-Milićević et al., 2017).

i) **Personalise the environment**: determines the nature of the online e-Learning environment.

ii) **Personalise the content**: Adopted content from the learners' personal environment.

iii) **Personalise the media**: Divergent content such as video or reading materials depending on learning styles and preferences.

iv) **Personalising learning sequences**: Allowing learners to choose their learning style by a nonlinear presentation.

v) **Personalise the conversation**: Using text or voice/video to adjust used sentences.

vi) **Personalise the navigation**: Permitting the learners to tour different parts of the content.

vii) **Personalise the learner**: Make the course personal to the learner.

viii) **Recognise individual competence**: Identifying different learners' needs and personalise accordingly.

ix) **Personalising learning objectives**: Assists learners to attain better learning objectives.

By harmonising the above-mentioned aspects, it is possible to achieve a truly personal learning environment for e-Learners which will facilitate the learners to learn according to their interest and even learn according to the preferred method of learning.

The above discussion regarding personalised Web search in e-Learning revealed some significant findings such as:

a) Popular traditional Web search engines (such as Google, Bing, Yahoo) serve as the top e-Learning tools for educational purposes among students. However, these systems show very limited potentials in offering a Personalised Web search result presentation for e-Learners.
b) Majority of the previous research works on e-Learning focused mainly on the personalisation of learning materials in different e-Learning applications. Very few studies (Capra et al., 2015; Curlango-Rosas et al., 2011) have noted how popular Web search engines can be used as an e-Learning tool in order to offer a personalised learning experience to students within a collaborative learning environment.
c) As far as we are concerned, so far no research work has been conducted by offering a Personalised Web search to students using the most popular Web search engines.

2.5 Recommender System

The explosive growth of the online environment has made it tougher for online users to search and select the information they are looking for from the Web. Users are overwhelmed by options (such as advanced search) which they might not have the knowledge or time to evaluate (Gavalas, Konstantopoulos, Mastakas & Pantziou, 2014). The recommender system (RS) is a proven tool which is used extensively in recent years for online users to cope with the information overload problem (Lu, Wu, Mao, Wang & Zhang, 2015). It can be characterised as a program that attempts to suggest the most appropriate items (suggestion of data or products or services) to particular users (businesses or individuals). Which is done by predicting the user's interest in an item based on relevant information about the items, the users and the interplays between the items and the users (Lu et al., 2015). The recommender system also aims to control the growing online information overload issue by suggesting the right items, services or source of information to users by applying data mining techniques and the prediction algorithms. The next section discusses the different types of recommendation techniques used by the recommender system to deliver recommendations.

2.5.1 General Architecture of Recommender System

A recommender system is comprised of four primary components (Figure 2.7) to provide recommendations followed by user profile, context, content repository and recommendation engine (Gavalas et al., 2014). These components are responsible for consuming information from relevant content repositories, elicit user requirements and capture situational context to deliver recommendations to users.

i) **User profile**: User profile is the most significant component of the recommender system. It accumulates the user's personal information such as preference, ratings, interaction with RS and user queries to send to the recommendation engine in order to generate recommendations.

ii) **Context**: Recommender system respect personal preferences and integrate many environmental and social contextual factors associated with the user through the context module concerning time and location. Subsequently, it passes the information into the recommendation generation process in order to provide a correct recommendation to a specific user.

iii) **Recommendation Engine:** Recommendation engine is accountable to gather information from other components of the recommender system. It processes the obtained information using various filtering techniques such as Content-based filtering (CBF), Collaborative filtering (CF) and Hybrid filtering techniques to create recommendations. In the next section, the different RS techniques will be elaborated briefly.

iv) **Content Repository**: The content repository stores the information relevant to the service that the recommender system offers. Recommendation engine selects contents from the content repository based on the user's interest to make recommendations.

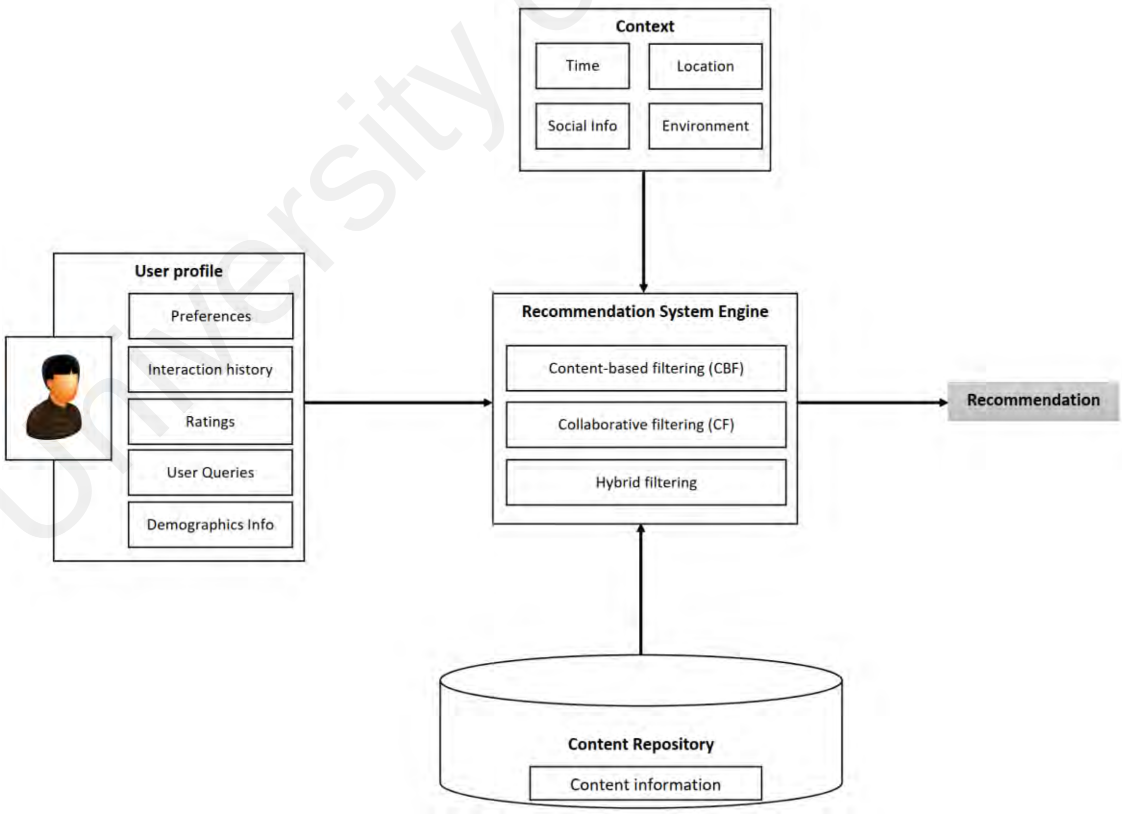

Figure 2.7: General architecture of the recommender system

2.5.2 Recommendation Techniques

To provide quality recommendations, it is very important for the recommender system to choose efficient and accurate recommendation techniques. This section summarises the most common and popular techniques used for recommendations.

i) **Content-based filtering (CBF) technique**: In this filtering technique, based on the user's profile a suggestion is made by employing the features derived from the content of the items, the user has assessed previously (Isinkaye, Folajimi & Ojokoh, 2015). The most relevant positively rated items are suggested to the user. The CBF method is solely based on the individual user's preferences. The CBF utilises diverse kinds of models to find similarities within documents in order to provide noteworthy recommendations. It may further use Probabilistic models such as Naı¨ve Bayes Classifier or Vector Space Model such as Term Frequency-Inverse Document Frequency (TF/IDF) (Islam, Wu, Ahmadi & Sid-Ahmed, 2007), Neural Networks (Introduction, 1994) or Decision Trees (Rokach & Maimon, 2010) to model the relationship between the different documents within a corpus. These techniques generate recommendations by understanding the underlying model with either statistical analysis or machine learning techniques.

ii) **Collaborative filtering (CF) technique**: Collaborative filtering is a domain-independent prediction technique. It works by building a database (user-item matrix) of preferences for items made by users. Then it matches users with related preferences and interests by measuring similarities within their profiles to deliver recommendations (Herlocker, Konstan, Terveen & Riedl, 2004). Such users create a group termed neighbourhood. A user gets recommendations to those items that he/she has not rated before but that were already positively rated by

Table 2.2: Key differences among recommendation techniques

Collaborative filtering	Content-based filtering	Hybrid filtering
▪ Collaborative filtering technique attempts to match users with similar interests based only upon queries or on suggestions or community connectedness. ▪ It focuses on recommendations but not on the level of ranking.	▪ Content-based filtering technique attempts to match based on the user profiles using features extracted from the content of the items the user has evaluated in the previously. ▪ It focuses on recommendations but not on the level of ranking.	▪ It combines different filtering techniques in order to generate recommendations. ▪ It focuses on recommendations but not on the level of ranking.

users in his/her neighbourhood. Recommendations that are generated by the CF can either be an estimation or a recommendation.

iii) **Hybrid filtering technique**: In order to gain a better system optimisation and evade some limitations and problems of simple recommendation systems, the Hybrid filtering technique is introduced. It combines different recommendation techniques (Stern, Herbrich & Graepel, 2009) to suppress the weaknesses of an individual technique in a combined model. Table 2.2 summarises the key differences among different recommendation techniques.

2.5.3 Personalised Recommendation for Web Search in e-Learning

Assisting learners in their learning when using Web search engines as an e-Learning tool is a significant research issue to address. In any learning context, students are presented with a huge number of results when they submit any search query. This causes information overload. They struggle to choose the right sources of content that will meet their learning needs, especially for the beginner level learners. Therefore, it is essential to leverage search engine users by integrating assistive components with current Web search engines that can help them find suitable learning contents effectively. In any learning context, it is also important to guide students in their Web search process. This can be achieved by providing these learners with a personalised recommendation that considers their individual profile differences. In the e-Learning context, it is even more crucial for the Web search engines to consider the students' academic background, their preferences, context and learning behaviour so as to be able to produce quality personalised recommendations which meet their learning needs.

2.6 Chapter Summary

This chapter has reviewed the literature relevant to this study. Firstly, an overview of the Web search engines is provided including personalisation in Web search. System components, requirements and techniques used in delivering personalised search results are also elaborated. Secondly, a brief discussion of e-Learning is provided along with the use of Web search engines in e-Learning. Several specialised Web search engines designed for e-Learners were also deliberated in terms of their key strengths and weaknesses. Thirdly, an overview of the recommender system, its general architecture, different types of techniques used for generating recommendations is described. It also included an outline on the personalised recommendation for Web search in e-Learning.

CHAPTER 3: METHODOLOGY

3.1 Introduction

This chapter illustrates the strategies used, further, details of the data collection process and approaches followed for the design, implementation and evaluation of the proposed system are discussed in this chapter which is followed by the documentation of the findings. A chapter summary is presented at the end of the chapter.

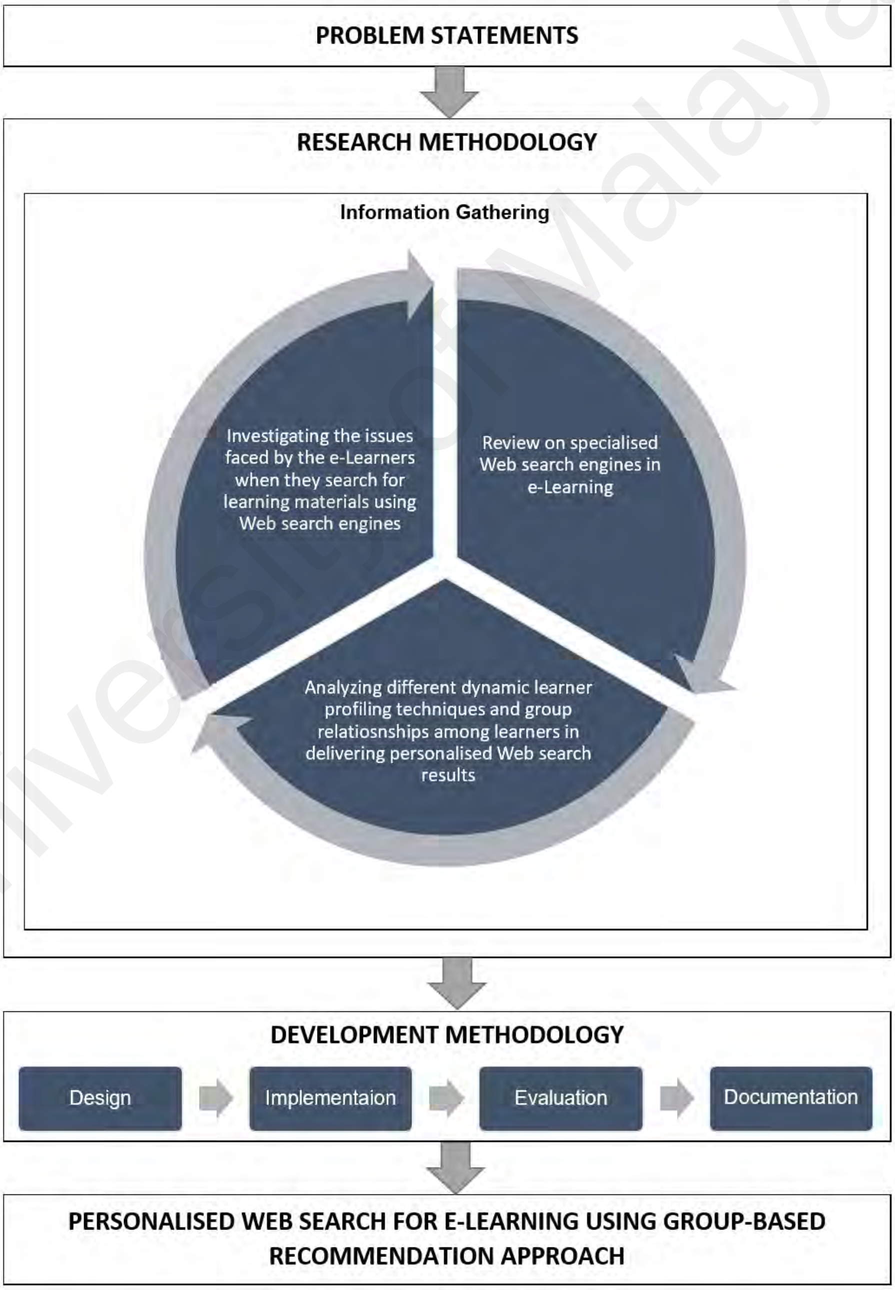

Figure 3.1: Research Methodology

3.2 Approaches to Research

Two types of methods are used in this study: research methodology and development methodology. The research methodology is depicted in Figure 3.1.

3.3 Research Methodology

The research methodology is the process of gathering information in order to find a solution to problems. The literature review was the primary method used in this study. It is used to collect, analyse and interpret information regarding issues faced by the students that limit their searching process when using Web search engines to search for e-Learning materials. It also performs studies on various existing Web search engines that are specially designed for e-Learners, including their strengths and weaknesses. Furthermore, an investigation is conducted in order to find the appropriate techniques that can be used to model the individual student's profile and to re-rank the search results to match the individual learning requirements. This is accomplished by providing personalised recommendations of Web search results. To conduct the literature review process, reputable scientific libraries, namely the IEEE Xplore Digital Library, ScienceDirect, SpringerLink, and the ACM Digital Library were intensively used to search for previous research works. Apart from that, search for relevant articles was also done via the Google Scholar, Microsoft Academic Search and Google Search Engine. In addition, the discussion approach turned out to be a useful technique to record important information besides the literature review with the supervisor during meetings. This also served as a guideline to complete this research work and to assure that the work was moving towards the right track.

3.4 Development Methodology

The development methodology refers to the process involved in developing the proposed system. It describes a common understanding of the activities involved in the development process. The development methodology of this study is comprised of three phases, namely: design, implementation and evaluation.

3.4.1 Design of the Proposed System

In this study, a framework of the proposed Personalised Web search for e-Learning system is designed for learners. The framework aims to provide users (students) with a personalised learning experience when using Web search engines. The framework offers students personalised search results as recommendations by using the group-based recommendation approach. In order to generate personalised recommendations based on the individual student's profile, the design phase will accomplish the following tasks:

i) **Student Profiling**: It is responsible for modelling student profile dynamically. It visualises personal data associated with specific users. The student profiling comprised two modules namely Academic Record Analyser and Behavioural Activity Record Analyser.

ii) **Content Selection and Re-ranking**: It analyses the shared interests of the similar group members on contents by investigating the relationships among the members in a group to re-rank the Web search results using Groupization technique and generates better personalised recommendations as search results.

The details of the design phase will be elaborated in Chapter 4.

3.4.2 Implementation of the Proposed System

The implementation phase is responsible for designing architectural model, database and user interaction of the proposed system. It also explains the implementation procedures of the system development and the technical requirements to deploy the system.

The proposed system is a Web-based system that follows three architectural layers namely: the application layer, the presentation layer, the database layer for implementation. The presentation layer is a Web portal which performs as a gateway between the students and the Web search engine (in this study Google search engine) via a Web browser (such as Firefox, Google Chrome, Internet Explorer, and Safari). The application layer consists of a Web search engine application which is integrated with a custom-built Web application employing the custom search engine API of Google. The database layer manages the physical storage and the retrieval of data. The application layer maintains an internal connection between the institutional Student management information system (MIS) and the local Web server's database repository through the database layer. The details of these implementation layers can be found in Chapter 5.

3.4.3 Evaluation of the Proposed System

The evaluation phase is accountable for assessing and evaluating the proposed method by testing the developed prototype for its ease of use, usefulness and effectiveness which is gathered from the experiment and questionnaires administered on the selected students.

A prototype was developed and hosted in the local server of the mentioned university. For the experimental purpose, the prototype was made available to the participants. From the second-year undergraduate students, 70 participants were invited who are studying in the Faculty of Computer Science and Information Technology, University of Malaya. Among them, 60 students participated in the experiment. Participants were then grouped into 4 groups, with each group consist of 15 randomly picked students. Total of four experimental sessions were conducted, each session belongs to a particular group. After completing each of the session, the following three sessions were also conducted with the successive three groups. Each of the group was then asked to accomplish a task which comprises of several problems represented by 25 multiple choice questions on advanced-level JAVA programming. Students were advised to use Google search engine via the developed prototype for helping themselves to find solutions.

Upon completion of the task, the students were requested to participate in a survey by completing a questionnaire. The aim is to find out their acceptance level after using the proposed system. Technology Acceptance Model (TAM) is used to formulate the questionnaire. TAM is developed by Davis (Davis, 1989), which is a tool often used to assess users' eagerness in accepting and using technology (Morris & Dillon, 1997). It has been widely used in several studies (Capra et al., 2007; Huang, Chien & Oyang, 2003). It consists of 12 questions for the purpose of this study. The first six questions assessed the users' perception towards the ease of use of our proposed system. The second set of six questions probed the participants' perception regarding how useful the system is. The questionnaire was presented using Google survey form which utilises a 5-point Likert scale with 1 meaning "completely agree" and 5 meaning "completely disagree". Value of the Cronbach's Alpha was calculated to measure the reliability of the scale, hence, the questionnaire set. In statistical analysis, Cronbach's Alpha is extensively used tool to assess the reliability of an instrument (Santos, 1999).

The one-way ANOVA and post-hoc tests were also performed to compare the results of each group which includes each profile level regarding the students' acceptance towards the proposed system and the search time used when solving the given problems. ANOVA test was applied to compare the averages of more than two groups (Cuevas, Febrero & Fraiman, 2004). The Tukey HSD post-hoc test was further used in this study to indicate which group would be significantly different from the others (Pignatiello, Camp & Rasar, 1986). A detail of the experimental evaluation result is presented in Chapter 6.

3.4.4 Documentations

After the completion of the testing & evaluation stage, as well as the other steps, the findings obtained from this study are documented in this dissertation. Moreover, the proposed approach and findings accomplished are also published in a conference proceeding and a journal paper.

3.5 Chapter Summary

The chapter discusses the research methodology used in this research. To attain the research objectives, a methodology framework is proposed. It comprises of several stages followed by information gathering, design, implementation, evaluation and documentation. The methods, instruments and techniques used in each of the phases are also explained in detail. Next chapter presents the framework design of the proposed system.

CHAPTER 4: FRAMEWORK DESIGN OF A PERSONALISED WEB SEARCH FOR E-LEARNING USING GROUP-BASED RECOMMENDATION APPROACH

4.1 Introduction

This chapter portrays the strategies and components used in the current study to design the proposed framework of a Personalised Web Search for E-Learning Using Group-based Recommendation Approach. First, it explains the purpose of the framework and the idea to form such a framework. Second, it explains the functionality of each component in detail. Finally, it elaborates on the detailed analysis of the techniques and approaches used to model the framework.

4.2 Overview of the Proposed Framework

A framework for the Personalised Web search for e-Learning using the Group-Based recommendation approach is proposed in this study. The purpose of the framework is to provide a middleware between the Google search engine and the institutional e-Learning portal that will deliver Personalised Web search results. The idea behind the formation of such a framework was motivated by the requirement to assist students in finding e-Learning materials that are related to their learning needs and their individual profiles. As pointed out by literature (Curlango-Rosas et al., 2011), many students prefer to use popular Web search engines to find resources that fulfil their learning activities. Unfortunately, when submitting a query, they are presented with a huge number of returned search results. This creates difficulties for them to select the right sources of content from the returned search results. Furthermore, the contents of the search results are the same for similar queries as the search engines do not take into account the individual's differences in learning. Such a situation impedes students' e-Learning activities, especially for novice learners who have even lesser experiences in dealing with web-based learning issues. It is worth mentioning that in YouTube, users get personalised video recommendations based on their personal profile match whereas in Google search engine, users do not receive any personalised recommendations. The only exception is perhaps there are some differences in result presentations due to the geographical location of users. For example, Google users from Australia will receive different arrangements of search result links compared with Google users from Malaysia. Therefore, it is very necessary to have such a framework that will leverage Web search engines, especially in the educational context so as to deliver personalised recommendations to students by analysing their learning needs and personal profiles.

Figure 4.1 shows the flow diagram of the online process of the proposed framework. The Web portal prompts secure login and authentication from the users. When a user tries to log into the system, the proposed system automatically identifies the users through the institution's Student Management Information System (Student MIS) and via local server records. When a student submits any query to the system, it is directed to the Google search engine as well as to the local server of the system. Consequently, a list of links is delivered by the Google search engine when certain keywords are used in the query. If the same query is used by every student, the returned links will be the same for all of them. However, the local server of the proposed system also returns a list of personalised links as recommended results based on the student's profile. In the profiling process, students are classified into three groups: Beginner, Intermediate and Master, based on their academic records and learning behaviours.

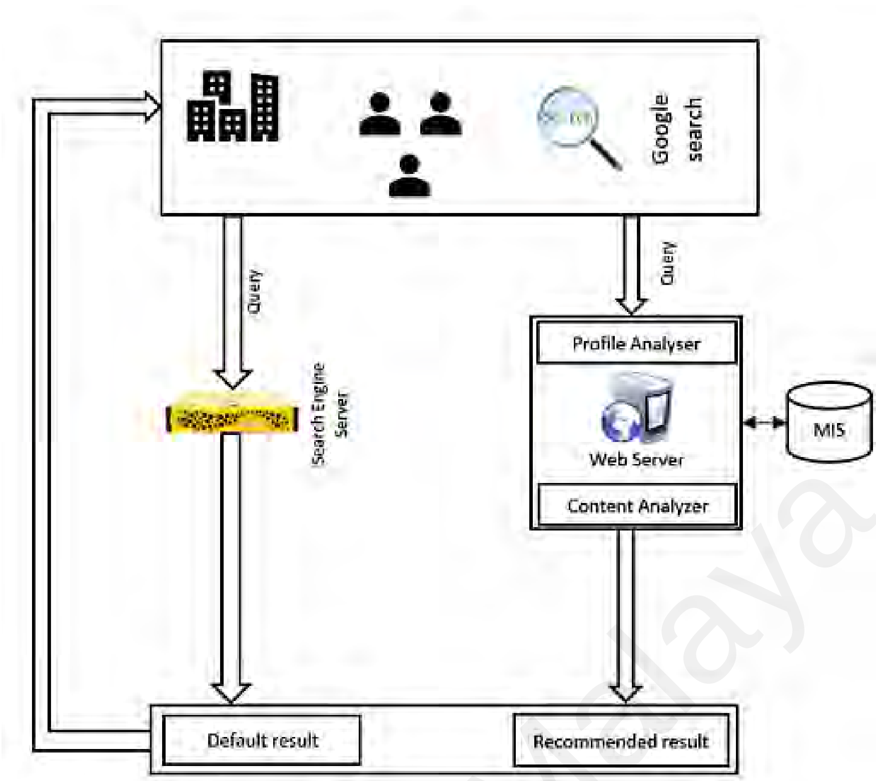

Figure 4.1: Online process diagram

When students search for the e-Learning materials using the Google search engine through the proposed system, their profiles are automatically updated, based on their learning behaviours. The behavioural data contain information regarding the students' Web browsing history and session logs when using the system. Any changes made in the academic records can also be detected by the system which then yields the updated information and dynamically updates the student's profile. Implicit feedback from students of each type of groups such as link selections, search activities and log histories are analysed and then stored in the local server. The system consistently compares this information with other students of similar profiles which results in a dynamic link recommendation outcome.

In order to update the students' profiles and content ranking list, a Web crawler is used. It identifies the changes in the students' academic records, learning behaviours and other related contextual records. The crawler is responsible for initiating the crawling process. It is a cyclical process. The crawler visits both the Student MIS and the local Web server within a 72-hour time span. It tries to fetch any changes that occur in the students' profiles and ranking lists. Any change manifested is sent to a profile analysing phase for keeping the students' profiles dynamically updated. The changes in the students' log histories and session data are detected in the profile analysing phase before the information is sent to a ranking analysing phase to be processed for re-ranking. The goal of the process is to produce the right set of recommendation links to the appropriate group of students. Next section will explain different components of the proposed framework including the process involved in each module to generate personalised recommendations in a Web search for students.

4.3 Student Profiling

The user profile is the key component in any e-Learning system (Premlatha et al., 2016; Wei & Yan, 2009; Yathongchai et al., 2013). It is a visualisation of the personal data associated with specific users (Wei & Yan, 2009). A profile refers to the explicit digital representation of a person's identity. The main goal of user profiling is customisation and the adaption of systems to suit users' specific needs (Fischer, 2001). In this study, student profiling is taken into account to create and maintain each individual student's profile. This will enable the system to understand the different learning needs and capabilities of each student. By using this information and by selecting the most relevant personalised links, the relevancy of the returned Web search results may be enhanced. This can be accomplished by prioritising the links according to each profile. To achieve this, the profiling process comprises two functional components: Academic Record Analyser and Behavioural Activity Analyser.

In this study, the students' profiles are constructed based on the IEEE PAPI standard (Lts, 2000; Wei & Yan, 2009) which is the most well-known user profiling model. The

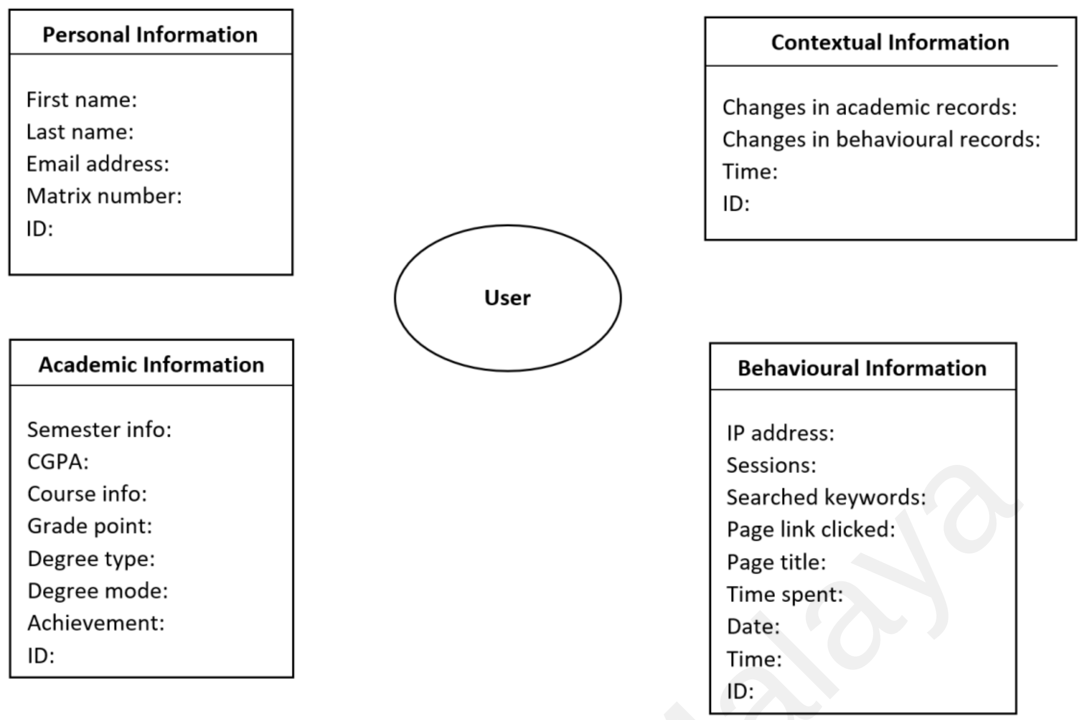

Figure 4.2: Student profiling model

model is further expanded by adding a learner profile design model (Wei & Yan, 2009). Figure 4.2 illustrates the model. Four types of data, Personal, Academic, Behavioural and Contextual records, are considered for modelling the students' profile. Several steps are involved in data collection. Firstly, a student's personal information is derived in order to identify the individual student. Such as first name, last name, email address and matrix number. Secondly, the students' academic records and their previous and current academic performance which are derived from the individual student's past and present academic performance are measured. These include semester number, cgpa. Course info, grade point, degree type, degree mode and achievement. Thirdly, the system inspects the learning behavioural activities of the students through their browsing histories and session data. Information such as user ID, IP address, sessions, searched keywords, page link clicked, page title, time spent and the number of logins is considered as behavioural activities. The purpose is to measure the students' interest level and their activeness towards learning while using the Web search engine. Whenever any change occurs in the students' academic and behavioural record, it is sensed and then exhibited in their respective profile updates. Finally, the system also discovers other coherent contexts (Guha, Gupta, Raghunathan & Srikant, 2015). This is achieved by tracking any change evidenced in the students' academic and behavioural records over a long time span and not just a few sessions or over a few days. However, the time span may range crossing months of data as contextual information begins to correspond to the dynamic profile needs. Figure 4.3 shows the processes involved in the student's profiling module.

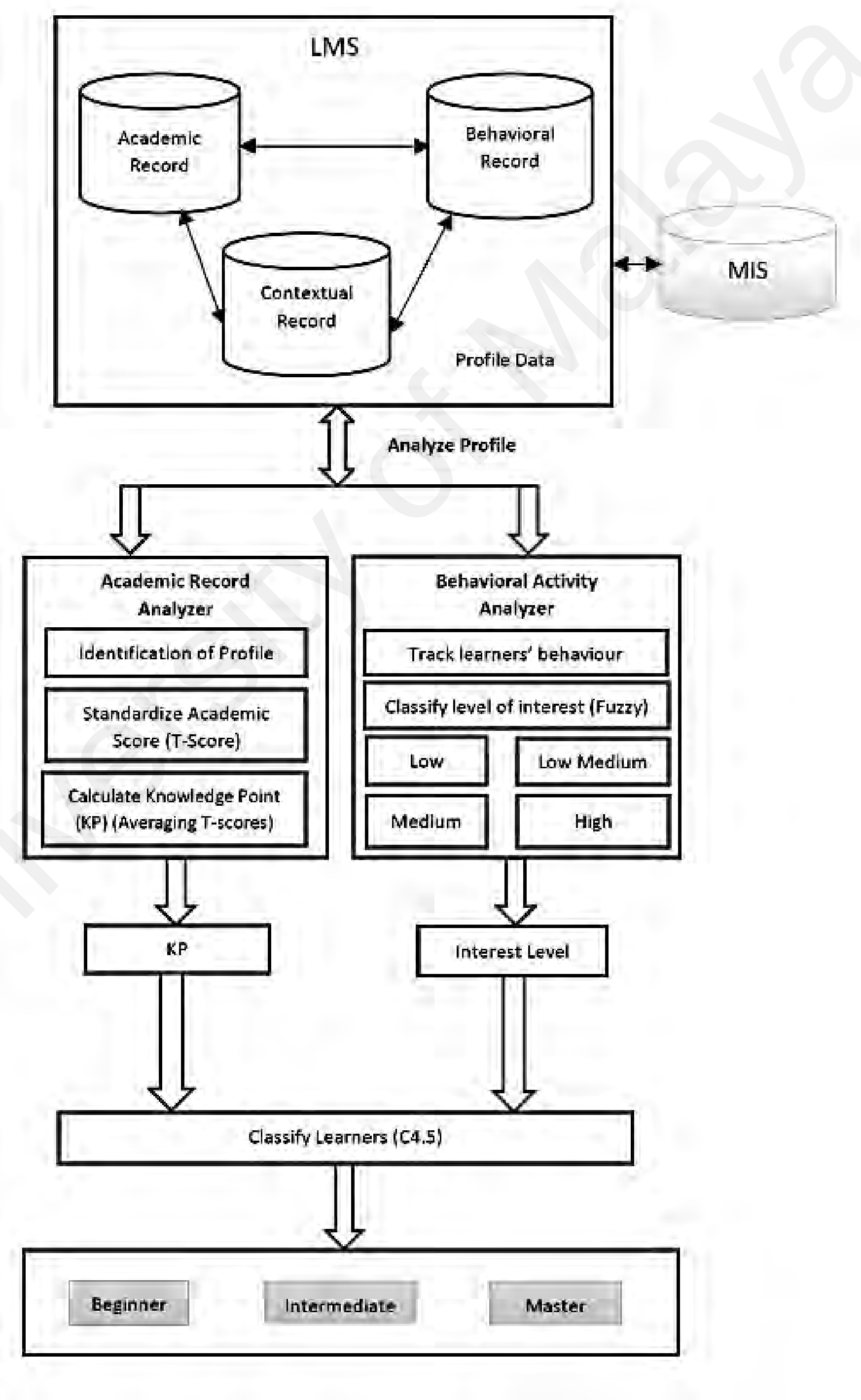

Figure 4.3: Student profiling module

4.3.1 Academic Record Analyzer

The primary function of the Academic Record Analyser is to identify the sign-in users and to retrieve their profiles from the Student MIS in the learner profile modelling process. The retrieved information mainly comprises the students' past and present academic records. Such information is then stored in the local server for further processing. Upon retrieval of the students' academic information, the module will calculate the standard T-scores for each student based on the raw scores of the academic records. This is achieved by using the grading policy recommended by the University of Texas at Austin (Iverson, 2017). The goal is to perform a comparison for each individual scores. T-scores is a form of standardised test statistics which provide the mean and standard deviation of a set of data (Carey & Delaney, 2010; Faulkner, Stetten & Miller, 1999; Krus & Krus, 1977). It generally brings about an improvement in all students' grades as compared to the approach which uses absolute percentages (Iverson, 2011). Alternatively, T-scores can also be used as benchmark scores since it eliminates the variation between grade points and can be used to decide whether the scores are high or low.

The subsequent step is to achieve an average score for each student by averaging the standard T-Scores. In this study, we define it as Knowledge Point (KP) for each student so as to classify the students' profiles. This allows an effective averaging of grades without the introduction of a bias in favour of tests with the greatest standard deviation. Since T-Scores are based on a normal (Gaussian) distribution, they generally represent the fairest way of grading. Nearly all national exams such as the SAT, MCAT, and GRE use a similar form of Standard T-Scores for evaluation (Iverson, 2017), thereby suggesting its effectiveness.

4.3.2 Behavioural Activity Analyser

The behaviour activity analyser module continuously monitors and captures the students' learning behaviours. This is achieved through their Web search activities while they use the proposed system. The learning activities extracted from their browsing histories and session logs are recorded and stored individually. The local server stores all the information and the captured data will contain information regarding the number of times they logged in, the number of searches done per login, the issued queries, the selected contents, page names, page sizes, links clicked, average scrolls, and time spent. In order to classify the students' level of interest towards learning, the stored activities are analysed while searching for the relevant learning materials from the Web search engine.

The students' learning behaviours, when viewed through the Web search, is often dynamic. A fuzzy setting can interpret their dynamic behaviours and interests. In this study, fuzzy rules are generated by applying the decision trees proposed by Sendhilkumar et al., 2014. The fuzzy rules classify the students' learning behaviours into four levels: Low, Low Medium, Medium and High. The classification also contemplates the pattern of the students' behaviours as representing the students' level of interest towards learning.

4.3.3 Student Profile Classification

The students' Knowledge Point (KP) and their level of interests have been mentioned in the earlier section 3.2.1. For the purpose of classifying every learner's profile adequately, further processing of these outputs is required. The students' academic performance (knowledge point) and learning behaviours (level of interest towards learning) are used to classify the students' profiles. This is achieved by using the extended classification rule (Premlatha et al., 2016; Yathongchai et al., 2013). A decision tree model is created by utilising the C4.5 algorithm (Quinlan, 1992) which is often referred to as the statistical classifier. The C4.5 can generate decision trees which are then used for inducing the classification models. Currently, it is the most powerful and preferred method in machine learning (Hssina, Merbouha, Ezzikouri & Erritali, 2014). As described by the creators of the Weka machine learning software, it is “a landmark decision tree program that is probably the machine learning workhouse most widely used in practice to date” (Witten, Frank, & Hall, 2011. p. 191). Its popularity appeared in the Top 10 algorithms for data mining (Sarker et al., 2015) and it is the most preferred and most powerful method to use (Hssina et al., 2014). The C4.5 classifier is considered in this study for the purpose of addressing the students’ profiles classification problem because of its supremacy, popularity and efficiency which are able to manage the delicacy of the student classification problem (Lakshmi, Indumathi & Ravi, 2016). Using Knowledge Point (KP) and the Interest Level, the decision tree (Figure 3.3) is translated into the following rules:

a) Students with a KP above 80 will be assigned the Master Class,

b) Students with a KP below 63 will be assigned the Beginner Class,

c) Students with a KP<=80, >=63 will be assigned depending on the following:

   (i) If his/her Interest Level is Low, he/she will be assigned the Beginner Class,

   (ii) If his/her Interest Level is Low Medium, he/she will be assigned the Immediate Class,

   (iii) If his/her Interest Level is Medium, he/she will be assigned the Intermediate Class,

   (iv) If his/her Interest Level is High, he/she will be assigned the Master Class.

This categorisation is further shown in Figure 4.4.

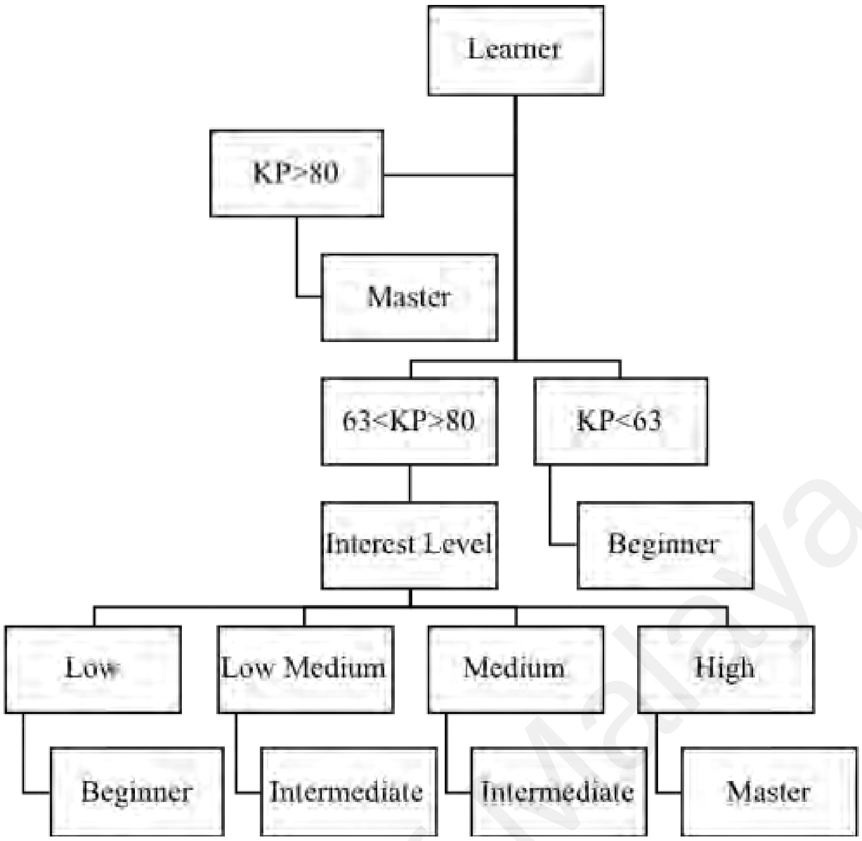

Figure 4.4: Decision tree for student profile classification

4.4 Content Selection and Re-ranking

A personalised Web search engine should have the ability to deliver results which consider user intent, needs and preferences (Leung et al., 2012; Leung et al., 2010). For a given query, a Personalised Web search engine is able to return different search results for dissimilar individuals. However, the collection of user data should be abundant enough for the system to understand the user's previous needs and preferences so that it can accomplish the personalised search outcomes, which is always a monumentally challenging task. Nonetheless, one way to overcome this is by combining the related data gathered from other individuals with similar profiles so as to build the personalisation features. In the "Groupization" technique, personalisation is established by putting higher weights on Web pages that appear more frequently in the Web history and the document terms of users from similar groups (Amer-Yahia et al., 2009; Amershi & Morris, 2008; Anuradha, 2012; Ben Ahmed et al., 2012; Teevan et al., 2009; White et al., 2013). This outcome can be achieved by harmonising with every group member's document term frequencies and Web histories. The requirement of the group recommendations is needed in many situations such as in recommending a movie for a group of friends to watch together (Said, Berkovsky & De Luca, 2011), a holiday destination for a family (Anagnostopoulos, Atassi, Becchetti, Fazzone & Silvestri, 2017), a restaurant for a group of co-workers to have lunch (Park, Park & Cho, 2008), or a set of e-Learning materials on a topic for a group of students (Rahman, Abdullah & Aurangozeb, 2017). Research (Brown, 2010; Rummel, Spada & Hauser, 2009; Watkins, 2009) has shown that students learn better in a well-structured cooperative group environment than in the traditional classroom set-up. Groupization is thus one of the methods which can improve the value of collaborative search tools in a group-based learning environment by using personalisations with shared interests. Shared interests of similar group members are taken into account so as to generate better personalisation results (document ranking level). However, the shared information is usually insignificant or unavailable to the search engines unless it encompasses additional information such as the type and length of the members' relationship among the search engine users (White et al., 2013). Since there is an inability of the available search engines to examine information regarding the relationships among the group members due to the lack of the users' public data and privacy concerns, our proposed Personalised Web search approach will be of advantage. Our proposed approach engages the group-based recommendation technique which takes advantage of the Groupization algorithm to discover different users' profiles within the homogeneous group of users (students) in an educational institution setting. This results in strengthening the approach by delivering more accurate personalised recommendations within the e-Learning domain.

In this study, we employ the Groupization technique in the content analysing phase so as to re-rank the returned search results into a custom sequence and to prioritise them in a way that is more suitable to the group members, depending on the members' similarity and level of preferences. Figure 4.5 is provided. The primary inducement behind utilising this Groupization technique is to improve the search results ranking, thereby making it more relevant according to the student's profile. To achieve this Groupization algorithm, group assessments on Web contents are performed. Priorities are given to the results which are considered to be useful to most of the members within the same category (similar profiles). In order to perform Groupization on a set of search results extracted from the Google search engine based on search query, firstly, a personalisation score is calculated for each result for every member of the group. Next, the Groupization score is calculated as the sum of the personalisation score of each group member.

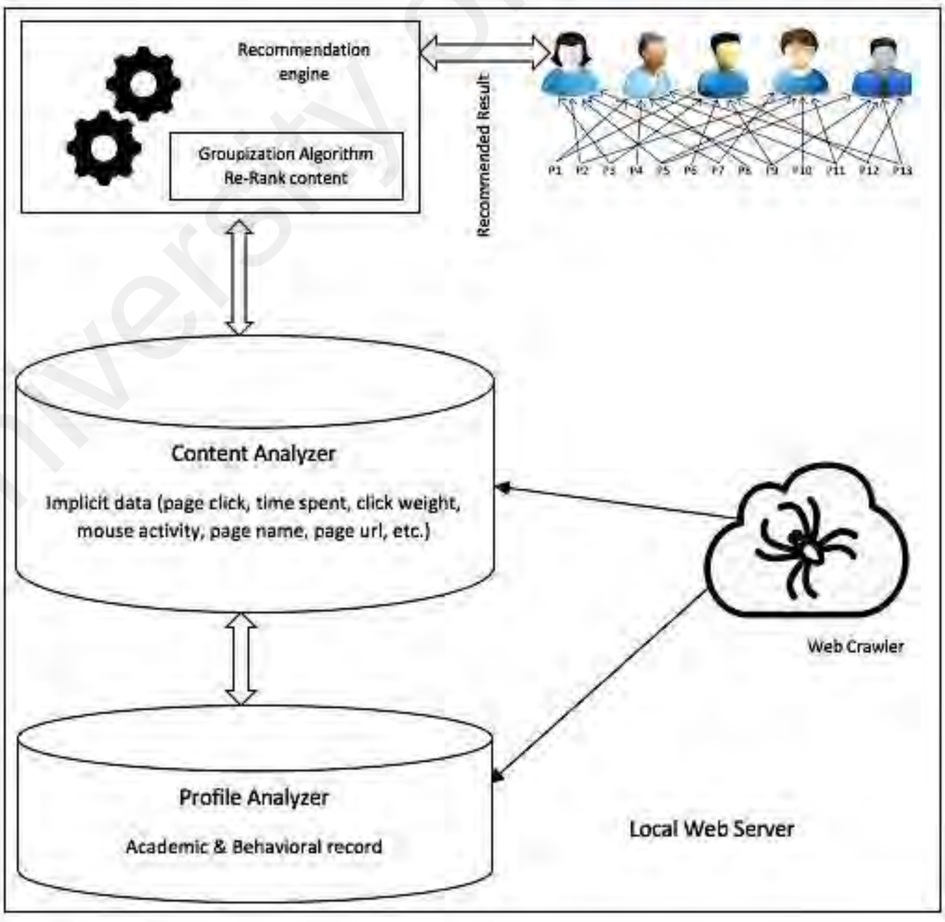

Figure 4.5: Content analyser (re-ranking module)

This is followed by a weighted combination of the Groupization scores and the original rank of the search results. The consideration of the latter is to preserve important information such as the results' "authoritativeness" (Ahmed et al., 2012; Ahmed, Nabli & Gargouri, 2013; Morris et al., 2008; Tamine & Pierre, 2016) used by the search engine. Let p1, p2, p3,…, pi (where i=1,2,3…,n) represent the set of pages returned by the search engine for a particular query, the weight of the page pi is given by

$$W_{pi} = \log \frac{(r_i + 0.5)(N - n_i + 0.5)}{(n_i + 0.5)(R - r_i + 0.5)} \tag{1}$$

where N is the total number of students in the sample, n_i is the total number of students who are in the various (Beginner/Intermediate/Master) levels, R is the number of students who visited page pi and who are at the Beginner/Intermediate/Master level. A higher page weight indicates a higher ranking. In the display section, only the top five most relevant recommended results are shown since research has indicated that students will spend more time on documents and learn better with the less number of returned result pages generated by the search engines (Kelly & Azzopardi, 2015). Different groups of students will receive different types of recommendation links. In the current study, the group of students are exemplified by the following characteristics:

- **Individual**: In this group, the queries, similar search tasks and link selections are obtained only from the current user's long-term history.
- **Group (Global)**: In this group, all the information is obtained from all the students' Web search history which focussed on similar tasks, queries, and link selections.
- **Group (Class)**: In this group, information regarding similar tasks, queries, and link selections is obtained from all the students within a particular group (i.e. Beginner, Intermediate or Master).

4.5 Chapter Summary

This chapter has outlined the theoretical framework used for the current study. It also outlined how the proposed framework working on the Personalised Web search system was developed for the current study and the reasons for using the respective steps were likewise indicated. The next chapter focusses on the implementation of the proposed system.

CHAPTER 5: IMPLEMENTATION OF THE PERSONALISED WEB SEARCH SYSTEM FOR E-LEARNING

5.1 Introduction

This chapter discusses the implementation of the Personalised Web search system for e-Learning by using the group-based recommendation approach. It explains the architectural design, the database design, the user interaction and interface design of the proposed system along with the technical requirements demanded in implementing and deploying the system.

5.2 System Architecture Design

The proposed personalised search system is a Web-based system that comprises three tiers of an architectural model, namely: the presentation layer, the architecture layer and the database layer. Figure 5.1 illustrates the architecture design of the prototype.

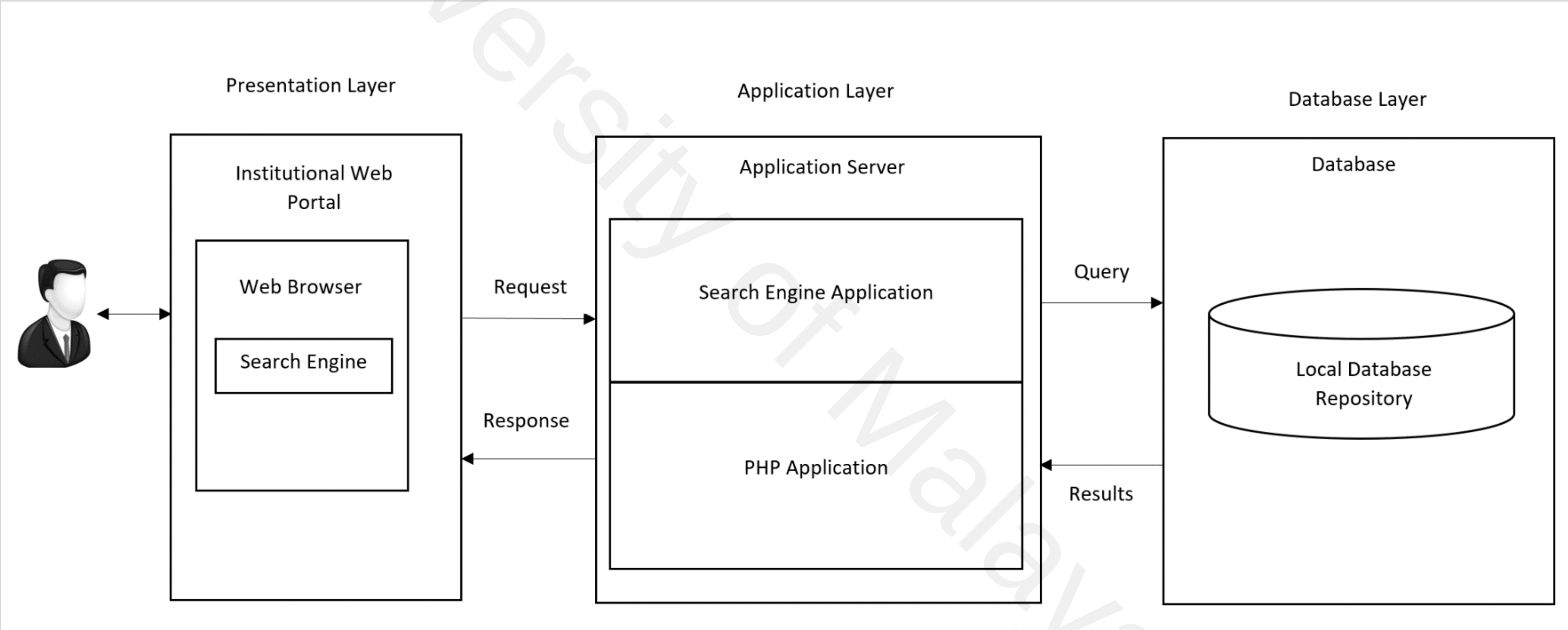

Figure 5.1: Personalised Web search system architecture design

The presentation layer is represented with an e-Learning Web portal which operates as a gateway between students and the Web search engine through a Web browser (such as Firefox, Internet Explorer, Google Chrome, and Safari). Through the gateway search engine, users (students) submit their queries and they receive results from the search engine. The Web portal is a custom Hypertext Preprocessor (PHP) built an application that integrates the public Web search engine (in this case Google Web search engine) through the Application Program Interface (API). Students are required to open the Web application using any Web browser, and the Web portal identifies the students through a secure login in order to use the search engine. They can perform any search tasks as usual as they are using traditional Web search engines. The presentation layer sends the user's request to the application layer to process the request. The application layer then passes the received request to the database layer and it then obtains the return results. Consequently, the presentation layer receives a response from the application layer and it then returns it to the user.

The application layer consists of the search engine application and a custom-built PHP application. The search engine application is then integrated with the developed Web application prototype using the API. The application layer receives an incoming request from the presentation layer, it handles all the system logic, manipulates data using data from the database layer and then interacts with the presentation layer to render the final output.

The database layer manages the physical storage and the retrieval of data. The application layer maintains a connection with the database repository through the database layer. It maintains an internal connection between the institutional Student MIS and the local Web server's database repository. The database layer hosts the students' academic record which is fetched from the institution's Student MIS and their learning behavioural information is then placed inside the local database repository. When students type any search query request to the Web search engine, the system retrieves a list of results from the search engine. The same request is also sent to the local Web server which also retrieves a list of personalised search results as recommendations. These two types of results are processed in the application layer after which, they are sent to the presentation layer to be displayed to the students.

The proposed system can be implemented in any educational institution's e-Learning infrastructure. A local Web server is required to host the PHP Web application and the database repository. The Web search engine is integrated by using API with the Web application. Students get the same search interface with similar features of the traditional Web search engine through our developed Web application. When a student types in any query, he/she receives a list of search results from the Web search engine along with a list of personalised search results as recommendations.

5.3 Database Design for the Personalised Web Search System for e-Learning

A database is a structured collection of data. To add, access, and process the data which are stored in a computer database, a database management system such as the MySQL Server is required. Since current computers are very good at handling large amounts of data, the database management systems play a crucial role in computing; they serve as standalone utilities, or as parts of other applications. There are different types of databases. The most common and popular one is the relational database management system (RDBMS). This system stores data in separate tables rather than putting all the data in one big storeroom. The database structures are organised into physical files and they are optimised for speed. The logical model, with objects such as databases, tables, views, rows, and columns, offers a flexible programming environment. Rules governing the relationships between different data fields such as one-to-one, one-to-many, unique, required or optional, and "pointers" between different tables can be setup. The database enforces these rules so that with a well-designed database, an application never suffers from inconsistent, duplicate, orphan, out-of-date, or missing data.

MySQL is the world's most popular open source database (Oracle, 2016). With its proven performance, reliability and ease-of-use, the MySQL has become the leading database choice for web-based applications. It is used by high profile web properties including Facebook, Twitter, YouTube, Yahoo! and many more. In the current study, we used the MySQL database to store the students' academic records and learning behavioural information which is located in the local server. As mentioned in the proposed system's architecture, the database server contains one database repository which stores the students' information such as their personal information, previous and present academics records, login information, Web browsing histories, and session information. The MySQL repository is structured using an Entity Relationship (ER) model which is one of the most widely used conceptual data models. An ER model represents data structure in terms of entities, relationships, and attributes (Moss, 2012). Figure 5.2 illustrates the conceptual design of the MySQL repository structure of the proposed systems.

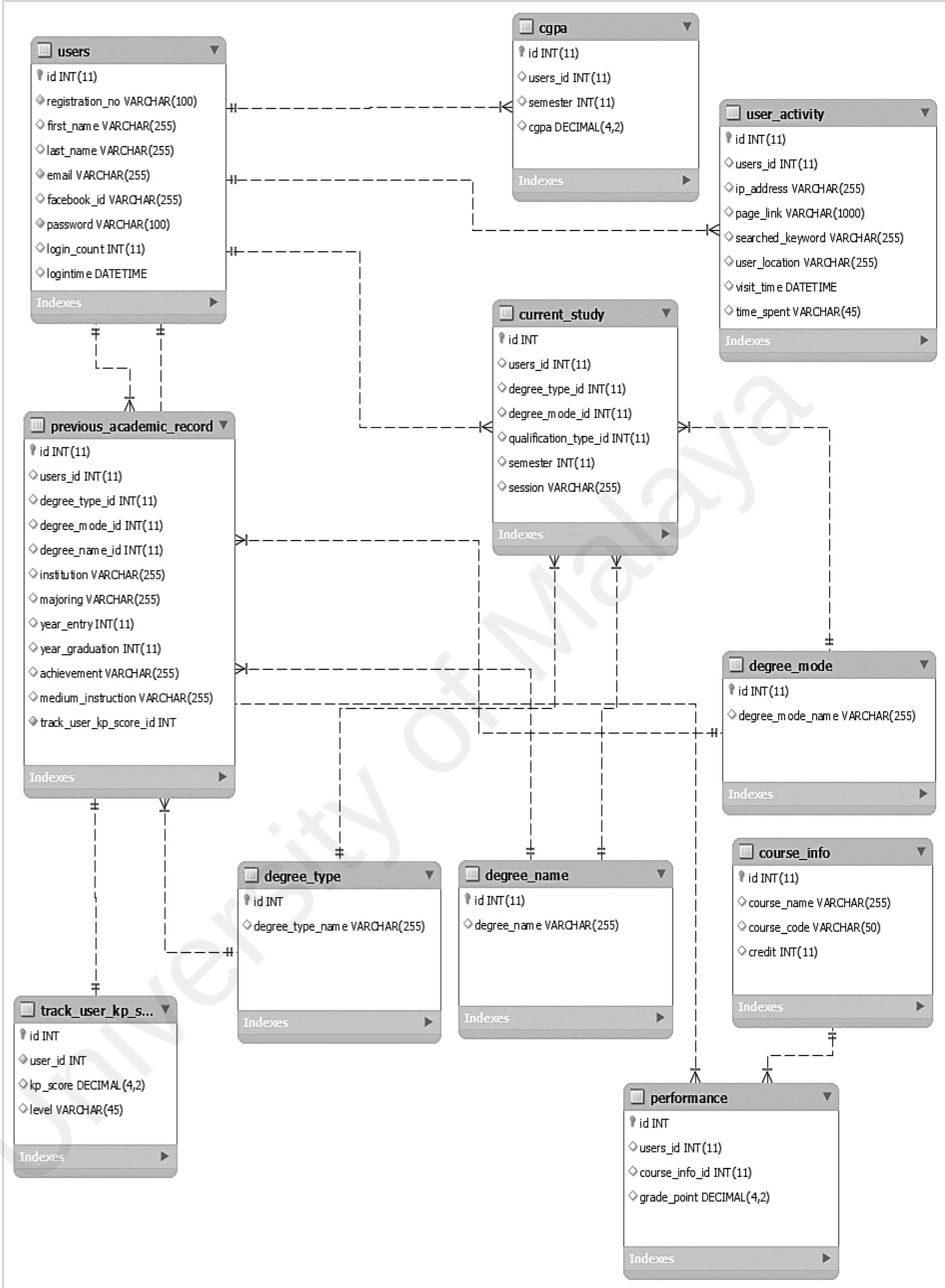

Figure 5.2: MySQL database structure

The above database structure is designed to represent a hierarchical data model which has one to many relationships. The structure presents information using parent/child relationships where each parent can have multiple children but each child can have only one parent. There are 11 relational tables in the Relational database management system (RDBMS) of the system. The table on ‘users’ stores all the information regarding the students’ personal identities and login information. The table on ‘previous_academic_record’ stores the learners’ past academic information including their degree, institution names, and achievement. The table on ‘current_study’ presents information regarding the students’ present academic records. Information related to the student’s achievement on courses is stored in the ‘performance’ table. The table on ‘cgpa’ stores information regarding the students’ overall Cumulative Grade point average (CGPA). The table on ‘course info’ collects information regarding the students’ academic courses with degree related information (such as mode, type, name) in ‘degree_mode’, ‘degree_type’ and ‘degree_name’ labels. The table on ‘user_activity’ stores the students’ learning behaviour while the table on ‘track_user KP_score’ stores the individual student’s calculated knowledge point and profile level. The details of each of the above-mentioned table’s structure such as field name, data type and description are shown in Tables 5.1 – 5.11.

Table 5.1: Database table ‘users’

Field Name	Data Type	Description
id	Integer	Identification number
registration_no	Variable Character Field	Student registration number (Matrix no.)
first_name	Variable Character Field	First name
last_name	Variable Character Field	Second name
email	Variable Character Field	Email address
password	Variable Character Field	Login password
facebook_id	Variable Character Field	Facebook identification number
login_count	Integer	Number of login
logintime	Timestamp	Login times

Table 5.2: Database table 'degree_type'

Field Name	Data Type	Description
id	Integer	Identification number
degree_type_name	Variable Character Field	Degree type

Table 5.3: Database table 'track_user_kp_score'

Field Name	Data Type	Description
id	Integer	Identification number
user_id	Integer	User identification number
kp_score	Decimal	Knowledge point
level	Variable Character Field	Learner's level

Table 5.4: Database table 'cgpa'

Field Name	Data Type	Description
id	Integer	Identification number
users_id	Integer	User identification number
semester	Integer	Semester Number
cgpa	Decimal	Cumulative Grade point average

Table 5.5: Database table 'course_info'

Field Name	Data Type	Description
Id	Integer	Identification number
course_name	Variable Character Field	Course title
course_code	Variable Character Field	Course code
Credit	Integer	Course credit hour

Table 5.6: Database table ‘current_study’

Field Name	Data Type	Description
Id	Integer	Identification number
users_id	Integer	User identification number
degree_type_id	Integer	Degree type identification number
degree_mode_id	Integer	Degree mode identification number
degree_name_id	Integer	Degree name identification number
Semester	Integer	Semester identification number
Session	Variable Character Field	Academic session

Table 5.7: Database table ‘degree_mode’

Field Name	Data Type	Description
Id	Integer	Identification number
degree_mode_name	Variable Character Field	Degree mode

Table 5.8: Database table ‘degree_name’

Field Name	Data Type	Description
Id	Integer	Identification number
degree_name	Variable Character Field	Degree title

Table 5.9: Database table ‘performance’

Field Name	Data Type	Description
Id	Integer	Identification number
users_id	Integer	User identification number
course_info_id	Integer	Course identification number
grade_point	Decimal	Grade point

Table 5.10: Database table 'previous_academic_record'

Field Name	Data Type	Description
Id	Integer	Identification number
users_id	Integer	User identification number
degree_type_id	Integer	Degree identification number
degree_mode_id	Integer	Degree mode identification number
degree_name_id	Integer	Degree name identification number
achievement	Decimal	Achieved grade point

Table 5.11: Database table 'track_user'

Field Name	Data Type	Description
id	Integer	Identification number
user_id	Integer	User identification number
ip_address	Variable Character Field	User IP address
start_time	Timestamp	Session start time
end_time	Timestamp	Session end time
searched_keyword	Variable Character Field	Searched keywords
page_link	Variable Character Field	Page link clicked
page_name	Variable Character Field	Page title
user_location	Variable Character Field	User location
short_desc	Variable Character Field	Page short description
is_recommendation	Variable Character Field	Recommendation clicked

5.4 User Interaction Design

This section describes the interface design of the Personalised Web search system for e-Learning by using the group-based recommendation approach. It is an important aspect for any system to provide a good searching interface to begin their Web search. The developed prototype developed by the current study has the same Web search interface design as Google's search engine. Figure 5.3 depicts the initial state of the Web search interface. The only difference is that it accommodates a section called 'recommendation for you' in the right section. In order to use the Google search engine through the proposed system, students need to log into the proposed system by using the same username and password as used in their institution's e-Learning portal. The login function verifies the authentic login and identifies the student's profile by using information extracted from the local server and the institution's Student MIS. After valid authentication, students are redirected to the Web search page (Figure 5.3).

Following the process, students can then type in the keywords and hit the search button in order to search for their desired contents. Students will then receive a list of personalised search results based on their identified learning profile, as indicated on the right side of the page. As usual, the default result list returned by the search engine is indicated on the left (Figure 5.4).

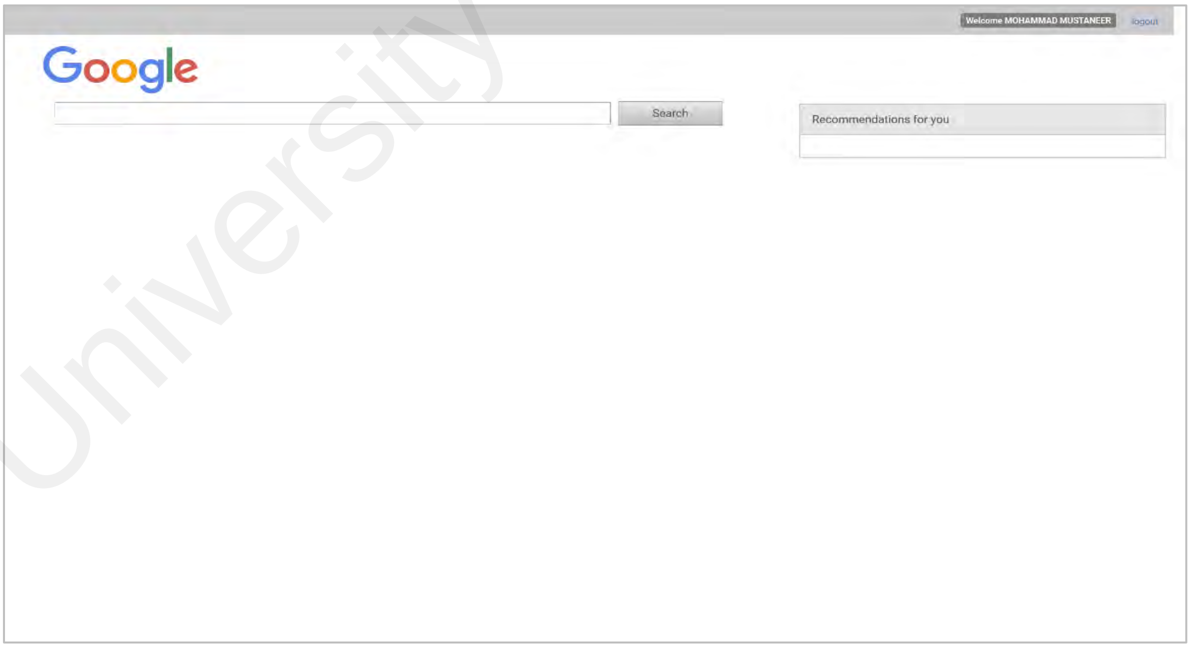

Figure 5.3: The initial state of our proposed Web search interface

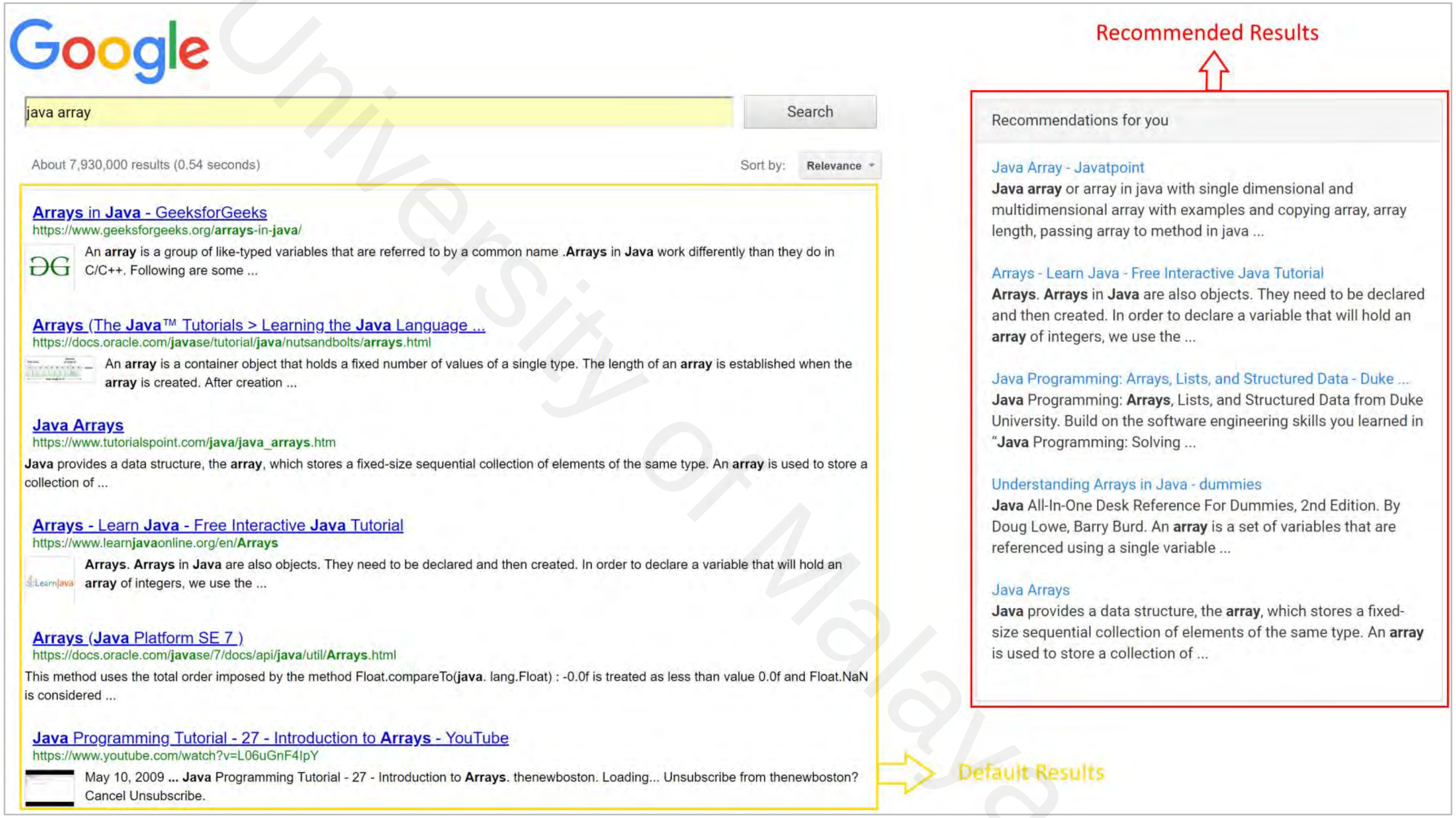

Figure 5.4: Interface of personalised Web search using group-based recommendation approach

5.5 Implementation

This section presents the implementation of the different modules involved in developing the Personalised Web search system by using the group-based recommendation approach. It also demonstrates the programming language used and the requirement of the software and hardware applied in developing the Personalised Web Search System for e-Learning.

5.5.1 Programming Language

The PHP programming language is used for developing the prototype. The PHP is a server scripting language; it is a powerful tool for making dynamic and interactive Web pages (Tatroe, MacIntyre & Lerdorf, 2013). It is widely used, is free, and is efficient for developing Web applications.

5.5.2 Development of the System

The CodeIgniter is a powerful PHP framework with a very small footprint. It is built for developers who can provide elegant toolkits to create full-featured web applications (Upton, 2007). It was created by EllisLab and it is presently, a project of the British Columbia Institute of Technology. It is an Open Source framework which has a very rich set of functionalities that may increase the speed of website development work. In the current study, the PHP CodeIgniter is used as an application development framework for developing the proposed prototype based on the Model-View-Controller (MVC) development pattern.

5.5.3 Implementation of the Academic Record Analyser Module

This module identifies the signed in users; it obtains their information from the Student MIS. It considers the students' previous and present academic records for student classification. It then calculates the standard T-Scores based on the extracted raw academic scores of each student, following the grading policy that was applied by the University of Texas at Austin. Consequently, an average score is achieved for each student by averaging the calculated standard T-scores in order to obtain the Knowledge Point (KP). The sample pseudocode is shown in Figure 5.5.

```
Function check academic records to calculate the Knowledge Point of students
        Pass In: student ID
        Get previous academic record to obtain students' previous academic achieve-
        ment
        Get current academic record to obtain students' present academic achievement
        from the current institution
        Get current cgpa of the students from the current institution
        Pass Out: result
End Function

Function calculate Knowledge Point (KP)
        Pass In: Students' academic records
        Initialize m to 10
        Initialize n to 77
        Calculate standard deviation of the previous academic records
        Calculate standard deviation of the current performance
        Calculate standard deviation of the cgpa
        Subtract sample mean of the previous academic records from the raw scores of
        the previous academic record, divided the result by standard deviation of the
        previous academic record and assign to academic score
        Multiply academic score by m and add n, and assign to t-score of the previous
        academic records
        Subtract sample mean of the current performance scores from the raw scores of
        the current performance, divided the result by standard deviation of the current
        performance scores and assign to current performance score
        Multiply current performance score by m and add n, and assign to t-score of the
        current performance
        Subtract sample mean of the cgpa scores from raw scores of the cgpa, divided
        the result by standard deviation of the cgpa scores and assign to cgpa score
        Multiply cgpa score by m and add n, and assign to t-score of the cgpa
        Calculate knowledge point (KP) of each student by averaging t-score of the
        previous academic record, t-score of the current performance and t-score of the
        cgpa
        Pass Out: result
End function
```

Figure 5.5: Sample pseudocode of academic record analyser

5.5.4 Implementation of the Behavioural Activity Analyser Module

The function of the module is to continuously monitor and capture the students' Web search activities in order to analyse their learning behaviours. This is achieved when students use the Web search engine through the proposed system. Each student's browsing histories and session logs are recorded and stored in the local Web server. The data captured comprises information such as login counts, page links, page names and so on. Here, students interest level towards learning is classified into four types: Low, Low Medium, Medium and High. The sample pseudo-code of tracking Web search behaviour while is shown in Figure 5.6. Figure 5.7 shows the pseudocode of interest level classification.

```
Function track student learning activities
        Pass In: student ID
        Track IP address
        Track location
        Track page link
        Track searched keyword
        Track page name of page clicked
        Track short description of page clicked
        Track click on recommendation links
        Track login count
        Track number of searches
        Track number of pages browsed
        Calculate search per login
        Calculate page size
        Store activities
        Pass Out: activities
End function
```

Figure 5.6: Sample pseudocode for tracking students' interest level

```
Function search per login calculation for fuzzification (average search per login)
        Pass In: Student ID, web search histories, session logs
        Initialize x value to 0, y value to 1, z value to 4, k value to 5
        if search per login is less than y
                set search per login fuzzy value one
        else if search per login is greater than x and less than k
                set search per login fuzzy value three
        else if search per login is greater than z
                set search per login fuzzy value four
        else error
        Pass Out: search per login fuzzy value
End function
Function Page size calculation for fuzzification (total number of browsed page)
        Pass In: Student ID, web search histories, session logs
        Initialize a value to 10, b value to 9, c value to 106, d value to 104, e value to 203
        if page size less than a
                set page size fuzzy value one
        else if page size greater than b and less than c
                set page size value two
        else if page size greater than d and less than e
                page size fuzzy value three
        else page size greater than e
                page size fuzzy value four
        Pass Out: page size fuzzy value
End Function
Function interest level classification
        Pass In: Student ID, search per login fuzzy value, page size value
        Initialize l value to 1, m value to 3, n value to 4, p value to 4, q value to 3, r value to
        2, s value to 1
        if search per login equal to 1 and page size fuzzy value equal to 1
                set interest level low
        else if search per login equal to l and page size fuzzy value equal to p
                set interest level low medium
        else if search per login equal to l and page size fuzzy value equal to r
                set interest level low
        else if search per login equal to l and page size fuzzy value equal to q
                set interest level low
        else if search per login equal to m and page size fuzzy value equal to s
                set interest level low
        else if search per login equal to m and page size fuzzy value equal to r
                set interest level medium
        else if search per login equal to m and page size fuzzy value equal to p
                set interest medium
        else if search per login equal to m and page size fuzzy value equal to q
                set interest level medium
        else if search per login equal to n and page size fuzzy value equal to p
                set interest level high
        else if search per login equal to n and page size fuzzy value equal to q
                set interest level medium
        else if search per login equal to n and page size fuzzy value equal to r
                set interest level medium
        else search per login equal to n and page size fuzzy value equal to s
                set interest level low medium
        Pass Out: interest level
End Function
```

Figure 5.7: Sample pseudocode of interest level calculation

5.5.5 Implementation of the Student Profile Classification

This classification module is responsible for classifying the learners into three groups based on their obtained knowledge point and their level of interest. They are categorised as Beginner, Intermediate and Master (as explained earlier). Each of the students belongs to any of these classes. The sample pseudo-code of student profile classification is shown in Figure 5.8.

```
Function student classification to classify students' profile
        Pass In: student ID, average T-Score, interest level
        set student level null
        if value of average t-score is greater than 80
        set student level master
        else if value of average t-score is greater than 63 and less than 83
                if interest level is low
                set student level beginner
                else if interest level is low medium
                set student level intermediate
                else if interest level is medium
                set student level intermediate
                else if interest level is high
                set student level master
                else set level error
        else set student level beginner
        Pass Out: student level
End function
```

Figure 5.8: Sample pseudocode for student profile classification

5.5.6 Implementation of the Content Analyser Module

In order to personalise the Web search results, the content analyser module re-ranks the search results based on the students' learning profiles. Different groups of learners receive different arrangements of the Web search results. The Groupization technique is responsible for re-ranking the search result based on the Web page weights. Higher page weight represents a higher page rank and page weight is determined by the popularity of the page based on the page hit by the students. The sample pseudo-code of the page weight calculation using the Groupization technique is shown in Figure 5.9. In addition, Figure

```
Function groupization to calculate the page weights of the search results
        Pass In: integer total number of students, integer total number of students
in a particular group, integer total number of students who visited a page, integer total
number of students who visited the page from a particular group
        Set N to total number of students
        Set ni to total number of students in a particular group
        Set R to total number of students who visited a page
        Set ri to total number of students who visited the page from a particular
group
        Initialize variable x to zero
        Initialize variable y to zero
        Initialize variable z to zero
        Initialize variable result to zero
        Initialize variable m to zero
        Initialize variable k to zero
        Add 0.5 to the ri and assign to k
        Add 0.5 to the ni and assign to m
        Subtract m from N and multiply with k and assign to x
        Subtract k from N and multiply with m and assign to y
        Divide x by y and assign to z
        Apply logarithm to z and assign to result
        Pass Out: result
End function
```

Figure 5.9: Sample pseudocode for the Groupization technique

```
Function page rerank to re order the page links suitable for different students' profiles
    Pass In: Integer user id, Integer search keyword, Integer total number of students
    in the sample, Integer total number of students in the sample of particular group,
    Character student's group
    Initialize variable total number of students who visited on a single page to zero
    Initialize variable total number of students of particular type who visited on a sin-
    gle page to zero
        While visited page counter is greater than zero for a search term
            Find page link
            Find page name
            Find description of pages
            Find total number of students who visited the page from a particu-
            lar group
            Call function groupization
            Pass total number of students in the sample of particular group,
            total number of students in the sample, total number of students
            who visited on a single page and total number of students of par-
            ticular type who visited on the single page to groupization func-
            tion as parameters
            Store results in the array
        end while
        Call Function build_sorter to short the array elements in a descending or-
        der for a searched keyword
        if total number of students who visited a page is greater than zero
            Print the Page name, Page link, Page description
    Pass Out: personalised recommendations
End Function
```

Figure 5.10: Sample pseudocode for re-ranking the Web page links

5.10 exhibits the pseudocode of the re-ranking of the Web search results which then provides recommendations to the students.

5.5.7 Integrating the Web Search Engine API

In this study, the Google Web search engine was used together with the proposed system in order to facilitate students in finding their desired learning contents. Students are requested to use the Google search engine in our developed prototype. Google API was used to integrate the search features with the prototype. The sample script for integrating the Google search engine using API is shown in Appendix A.

5.5.8 Making Connections to the Database

In order to be connected to the MySQL database server, we used the CodeIgniter database library class. The database model corresponds to all the data-related logic that the student works with. This also represents the data that are being transferred between the View, Controller components (user controller) and the logic-related data. The sample script of connecting with the database is shown in Appendix A.

5.5.9 Creating the SQL Operational Statements

Once a connection is established, interaction with the database is granted. This involves the create, read, update and delete operations process. Several SQL statements were generated when using the CodeIgniter database class to perform the various tasks such as adding the learning behavioural activities and the student records, updating the knowledge points and others. The sample script of some of the operational SQL statements is shown in Appendix A.

Table 5.12: Software Requirements

Tools	Software
Operating System	Windows 8
Web Server	Linux
Database Server	MySQL (5.1+)
Server-Side Scripting Language	PHP 5.6.37
Object-Oriented Framework	PHP CodeIgniter Web Framework
Client-Side Scripting Language	jQuery
Tools	Notepad++, MySQL work bench
Web Browser	IE 6.0 or above, Mozilla Firefox, Chrome

5.5.10 Software and Hardware Requirements

There are several software and hardware requirements required in order to develop the proposed Personalised Web Search system for e-Learning when using the group-based recommendation approach. The requirements needed for the essential system are illustrated in Table 5.12 which depicts the minimum specifications needed to run the system.

5.6 Adopting the Framework

The proposed system can be adopted by any group-based e-Learning activity involving the process of searching for Web materials within an educational institution environment. In order to integrate the proposed group-based e-Learning system using Web search, it is necessary to have a computer server, access to the Student MIS as well as access to the Web search engine via the institution's internet. Below are the guidelines developed for people who intend to establish the proposed collaborative e-Learning environment for their students.

Step 1: Integrate the proposed framework within the institution's e-Learning portal.

Step 2: Setup the Web and database server.

Step 3: Create individual login access for each student.

Step 4: Request students to use the Web search engine via the proposed system.

Step 5: Motivate students to use the proposed system within the institution for educational purposes only.

5.7 Chapter Summary

This chapter has described how the prototype system was developed for the current study. It has proposed various steps which can enable students using the Web search engine to look for learning materials to be more familiar with the system. This chapter has also explained the various requirements needed to develop the personalised Web search system for e-Learning.

CHAPTER 6: EVALUATION AND DISCUSSION

6.1 Introduction

After the implementation of the proposed Personalised Web search system for e-Learning using the group-based recommendation approach, the prototype needs to be evaluated for its effectiveness. The evaluation has to play a part in order to appraise the ease of use, its usefulness and its effectiveness. This chapter discusses the experimental design, procedure, materials used, data collection and evaluation method for evaluating the proposed system. The result of the evaluation is analysed and presented in the last section of this chapter.

6.2 Experimental Design

In order to ensure that the proposed Personalised Web search system for e-Learning using the group-based recommendation approach fulfils its requirement, an experiment was conducted. A group of second-year undergraduate students from the Faculty of Computer Science and Information Technology, University of Malaya, was recruited. A prototype was developed and then hosted in the said university's local server. The Google search engine was integrated by using Google API. For the purpose of the experiments, the prototype was made available for the participants to access the system by creating individual accounts. Each participant was required to log into the system from the login page in order to use the Google search engine. Participants were then asked to solve the given problems with the help of the Google search engine.

6.2.1 Participants

The sample in this experiment consists of 60 second-year undergraduate students. They were among the 70 students who were invited to participate in the experiment. These 60 students were then grouped into 4 groups, with each group consisting of 15 randomly picked students. The purpose of this experiment was to measure the effectiveness of our proposed Personalised Web search recommendation system, particularly its ability to assist students in searching for e-Learning materials within a group-based learning environment. The intention of selecting students with a background in Computer Science is that they were assumed to be more technically sound in utilising Web search engines.

6.2.2 Materials

In this experiment, the participants were given 25 multiple choice questions on advanced-level JAVA programming. Table 6.1 lists the questions used for the experiment. Participants were required to answer all the questions within 45 minutes. The advanced-level problems were given because it had been reported in the literature that Web search engine users tend to have more interactions with search engines while solving complex tasks as compared to simpler ones (Capra et al., 2015). This is important because one of the key factors in this experiment was to ensure that the participants would rely on the Web search engine to solve the given tasks as much as possible.

Upon completion of the task, the participants were requested to complete a questionnaire. The questionnaire would determine their level of acceptance of the proposed system. The questionnaire was adapted from the Technology Acceptance Model (TAM). It consists of 12 questions for the purpose of this study. The first six questions assessed the users' perception towards the ease of use of our proposed system. This is shown in Table 6.2

Table 6.1: Experimental questions on java programming

No	Questions
1	What do you mean by platform independence in java?
2	What is the difference between a JDK and a JVM?
3	Which arithmetic operations can result in the throwing of an ArithmeticException?
4	What is the difference between Path and Classpath?
5	When can parseInt() method be used?
6	java.util.regex belongs to which class?
7	Which two methods would you need to implement key Object in HashMap?
8	What is the difference between TreeSet and SortedSet?
9	What is an immutable object? Can you create an immutable object?
10	What is the difference between StringBuffer and StringBuilder in Java?
11	What will be the problem if you don't override the hashcode() method?
12	What is the difference between CyclicBarrier and CountDownLatch in Java?
13	What is applet Lifecycle?
14	What is Externalizable?
15	What is constructor chaining and how is it achieved in Java?
16	What is Downcasting?
17	What is a transient variable?
18	Define JIT compiler.
19	What is the difference between ArrayList and Vector?
20	What is busy spin?
21	Is Swing thread-safe?
22	What is an immutable object? How do you create an Immutable object in Java?
23	Mention one difference between WeakReference and SoftReference in Java?
24	What is the size of int in 64-bit JVM?
25	What is constructor chaining in Java?

The second set of six questions probed the participants' perception regarding the usefulness of the system, as shown in Table 6.2. Here, the ease of use is defined as the degree to which the user believes that the system would be free of effort whereas usefulness is defined as the degree to which the user believes the system would improve his/her task performance. The questionnaire was presented in the form of a Google survey form which utilises a 5-point Likert scale with 1 meaning "completely agree" and 5 meaning "completely disagree".

Table 6.2: Questionnaire based on TAM Model

Participants' perceptions of ease of use	
Q-01	Using the system to find learning materials (answers) is easy for me.
Q-02	I would find it easy to get the system to do what I want it to do.
Q-03	My interaction with the system is clear and understandable.
Q-04	I find the system easy to use to look for learning materials (answers).
Q-05	I find the system flexible to interact with.
Q-06	It is easy for me to become skilful at using the system.
Participants' perceptions of usefulness	
Q-07	Using the system to look for learning materials would enable me to accomplish search tasks more quickly.
Q-08	Using the system to look for learning materials would improve my e-learning.
Q-09	Using the system in my e-learning would improve my search productivity.
Q-10	Using the system to look for learning materials in my e-learning would enhance my effectiveness.
Q-11	Using the system to look for learning materials would make it easier to do my e-learning.
Q-12	I find the system useful in my e-learning environment.

6.3 Evaluation Procedure

The experiment comprised four sessions. One session was assigned to each group. The first session began by getting the 15 participants to solve all the 25 problems on advanced-level JAVA programming within 45 minutes. The only restriction imposed on the participants was that they had to search for the solutions by using the Google search engine via our developed prototype. This was ensured by recording the browsing history, session logs and completion time. At the end of 45 minutes upon completion, the participants were asked to respond to the questionnaire on the Google survey form which probed their perceptions towards the ease of use and the usefulness of the proposed system. Upon the accomplishment of the first session with the first group, the next three sessions were conducted with the other three groups, successively. The same set of problems and questionnaires were given to all these three groups in all the sessions.

6.4 Data Collection

The data collection phase was made up of two phases. The first phase involved fetching the participants' personal and academic information from the Student MIS and storing the information in a local server. The second phase involved capturing the participants' learning behaviour while using the proposed system. This is achieved by browsing their history and session data where the data of every participant were automatically stored in the local server system while using the prototype.

6.5 Evaluation Matrix

In order to determine the students' acceptance towards the proposed Personalised Web search system for e-Learning using the group-based recommendation approach, a survey was conducted through a set of questionnaire which was adopted by the Technology Acceptance Model (TAM). In addition, to measure the scale of the reliability of the questionnaire set, Cronbach's Alpha test was also performed.

In this study, we compared the results of each group with regard to the students' acceptance towards the proposed system and the search time spent when solving the given problems. To achieve this, the one-way ANOVA and post-hoc tests were performed. Furthermore, to indicate which group would be significantly different from the others, The Tukey HSD post-hoc test was also used.

6.6 Evaluation Results, Analysis and Findings

This section looks at the analysis of the results gained from the experiment.

6.6.1 Analysis of the Development Outcome

The increasing diversity and richness of the contents in the Web call for the need to revise the concept of presenting the same search results for all users. Although notable personalised search applications are currently available to assist users in using the Web as a resource, most are catered to areas not related to learning, for instance, travelling, shopping, entertainment and personal needs only. To the best of our knowledge, a similar system for the purpose of helping users to find personalised e-Learning materials using the Web search engine is still non-existent. Furthermore, we believe that the students are those who urgently require personalised recommendations that can enhance their e-Learning materials because, among learners, there are learner differences. To bridge this gap, a novel Personalised Web Search system for e-Learning was thus created and this was established from the group-based recommendation approach. Our prototype aims to leverage students by delivering more personalised search result recommendations that match their individual profiles.

In order to evaluate the reliability and acceptability of the proposed system, we then conducted an experiment on 60 undergraduate students. From this experiment, the users' perception of the ease of use and the usefulness of the proposed system was probed. We further obtained additional information from the experiment by focussing on the students' search performance and the time taken to achieve their tasks while using the prototype. However, the accuracy of the search results such as evaluation of precision and recalls were excluded from this study.

In the experiment, the participants were given a set of problems on advanced-level JAVA programming questions. They were required to find the solutions using the Google search engine via our prototype. As shown in Figure 6.1, the proposed approach returns personalised search results that matched the students' queries (such as "TreeSet and SortedSet") as well as their personal profiles (such as Beginner, Intermediate and Master). The left side of Figure 6.1 shows the typical search results returned by the Google search engine which are the same for all users with the same query. In contrast, the right side of Figure 6.1 shows the recommended links returned by the proposed system. It can be observed that the recommended links provided by the prototype were dissimilar for different types of student profiles. Each profile belongs to one of the three groups: Beginner, Intermediate and Master.

The proposed system provides students with the top five recommended links which are personalised based on the group that they belong to. For example, students from the Intermediate and Master group will get different recommended links for the same query, as shown in Figure 6.2 and 6.3, respectively. The order of the links is sorted based on the most relevant weighted page. For instance, the link titled 'Java Collection Tutorial – Java Sorted Set' appears 2nd in the order for the Intermediate group whereas it is unavailable for the Master group. This is due to a higher weight being assigned by the 'Groupization' algorithm to the link as the Web page was preferred by many members of the Intermediate group. However, most of the students from the Master group did not prefer the page, consequently, a lower page weight was assigned. It was very much lower than the weights of the top five links observed in the recommended list for the Master profile. In another instance, the 'set – the difference between navigableSet, SortedSet and TreeSet in Java? - Stack Overflow' appears in both profiles. This is because most of the students from the Intermediate and Master groups considered the page to be relevant in searching for their answers. Therefore, a higher weight was assigned by the 'Groupization' algorithm for the link for both profiles. The weight was determined by the total number of hits made by the students within the same group in descending order. While students used the system, the group was determined by their academic record assessments and their learning behaviours, as detected by the system. As a consequence, different sets of recommended links will be presented to different groups of students, depending on the different arrangement of priorities.

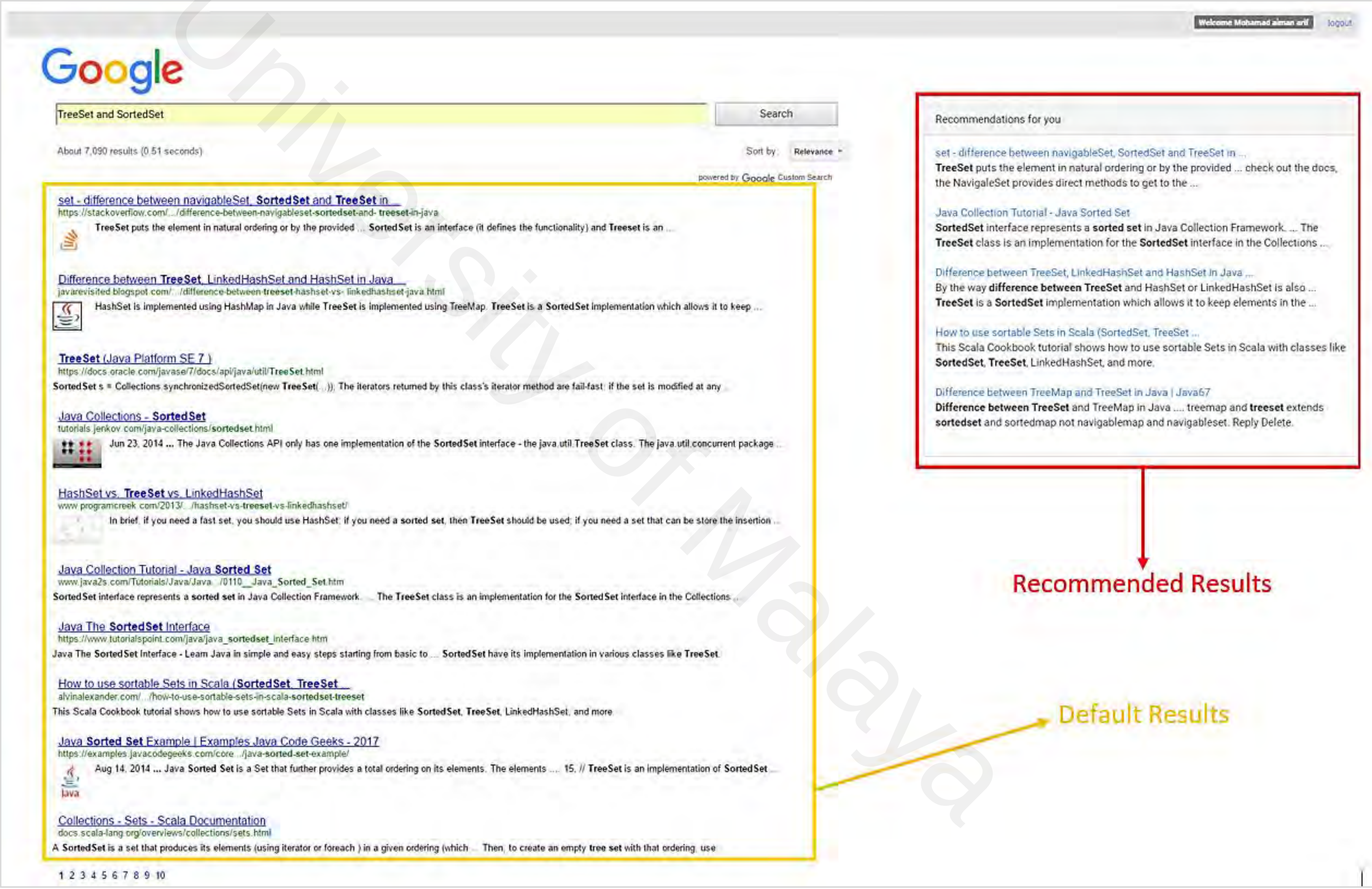

Figure 6.1: Search Interface and recommendations list display

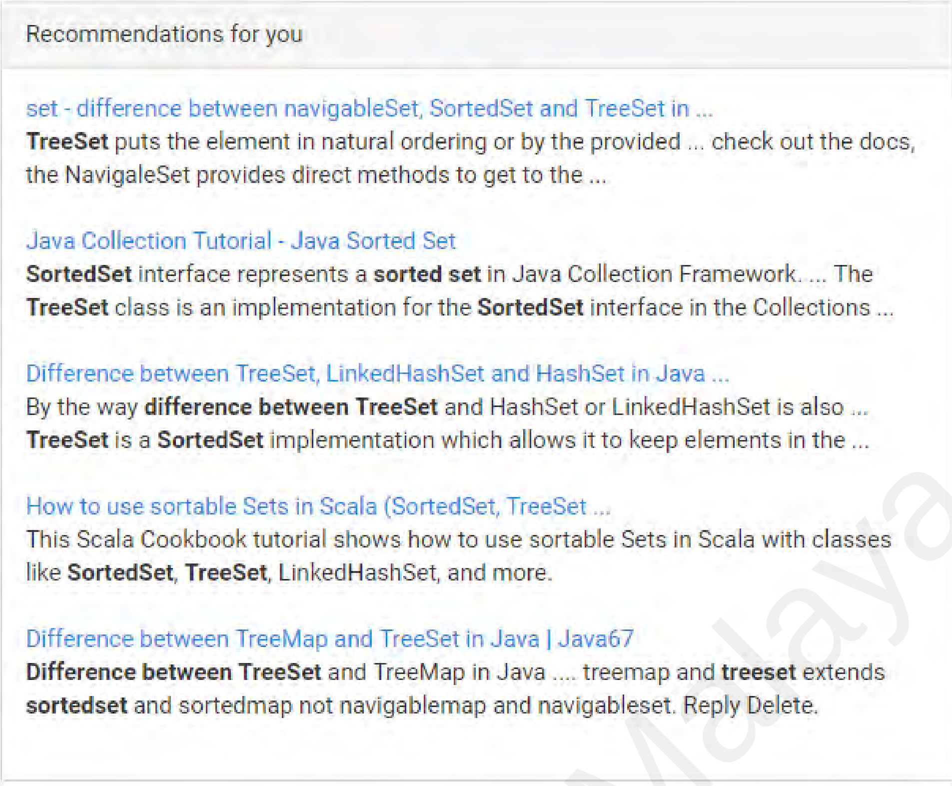

Figure 6.2: Display for the Intermediate profile

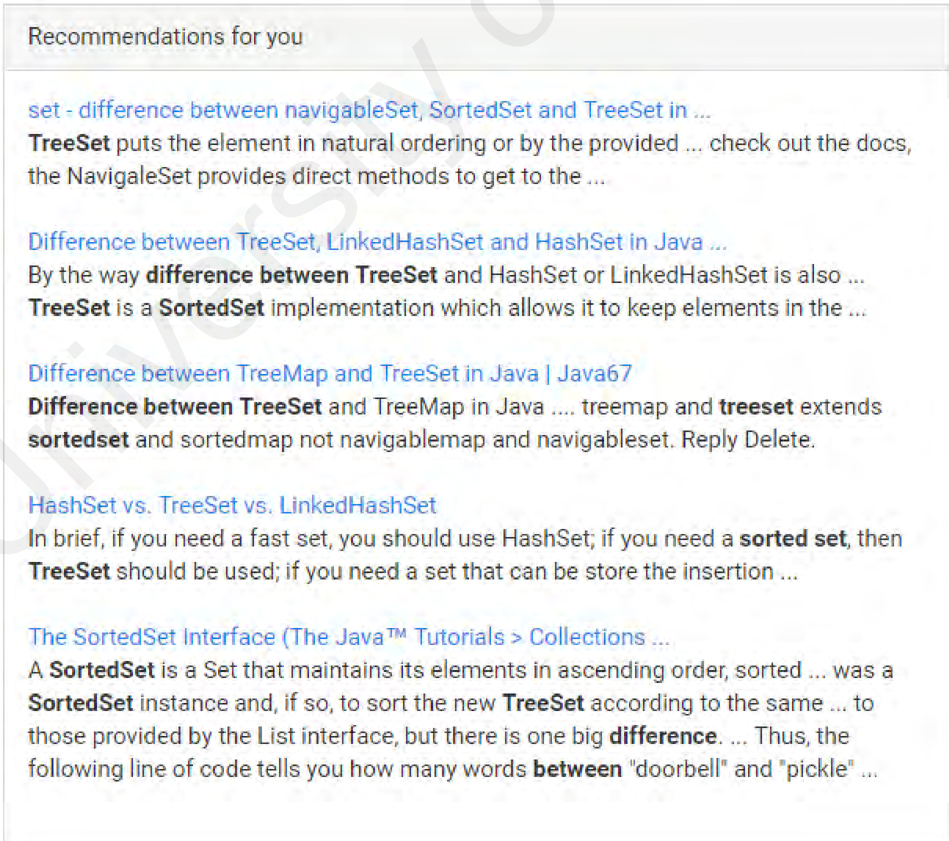

Figure 6.3: Display for the Master profile

6.6.2 Analysis of Student Acceptance

With the aim of assessing the students' level of acceptance towards our proposed Personalised Web search system for e-Leaning using the group-based recommendation approach, we also conducted a survey. As mentioned, the TAM questionnaire was applied to determine participants' "perceived ease of use" and "perceived usefulness" towards the developed prototype, particularly, their perception in finding relevant resources for solving the problems. The reliability of the questionnaire had been assessed and the Cronbach's Alpha test showed a high-reliability score of around 95%.

Table 6.3: ANOVA results of the four experimental groups (TAM questionnaire)

Scale	Groups	N	Mean	SD	F	Post hoc tests	CI
Ease of use	Group 1 (a)	15	3.57	0.95	6.11	a < b	95
	Group 2 (b)	15	3.91	0.78		b < c	
	Group 3 (c)	15	4.26	0.63		c < d	
	Group 4 (d)	15	4.66	0.49			
Usefulness	Group 1 (a)	15	3.72	0.86	4.19	a < b	95
	Group 2 (b)	15	3.93	0.76		b < c	
	Group 3 (c)	15	4.34	0.63		c < d	
	Group 4 (d)	15	4.53	0.50			

*p<0.05.

Table 6.4: ANOVA results of the three profile levels (TAM questionnaire)

Scale	Groups	N	Mean	SD	F	Post hoc tests	CI
Ease of use	Beginner (a)	13	4.03	0.76	0.0264	a >b	95
	Intermediate (b)	26	4.02	0.87		b < c	
	Master (c)	21	4.07	0.60			
Usefulness	Beginner (a)	13	4.11	0.82	0.0937	a > b	95
	Intermediate (b)	26	4.01	0.71		b < c	
	Master (c)	21	4.02	0.63			

*p<1.

The ANOVA results generated by the TAM questionnaire for all the four experimental groups are presented in Table 6.3. The outcome of the one-way ANOVA test and the post-hoc test of the four groups of participants are similarly included in Table 6.3. The results showed that the average ratings for perceived "ease of use" ($F=6.11$, $p<0.05$) were 3.57, 3.91, 4.26 and 4.66 for Group 1, 2, 3 and 4, respectively. In comparison, the average ratings for "perceived usefulness" were 3.72, 3.93, 4.34, 4.53 for each of the four groups respectively. These results denote that there were significant differences between the means for all the four groups in both dimensions. It also indicated that Group 1 had the lowest acceptance whereas Group 4 had the highest acceptance towards our proposed system in terms of given ratings. Group 4 had the highest degree of satisfaction in terms of ease of use and usefulness, followed by Group 3, 2 and 1. This implies that Group 4 and the students who participated in the sessions had revealed the highest degree of satisfaction, comparatively.

When Group 4 was requested to find solutions to the given problems, the proposed Personalised Web Search system using group-based recommendation approach enabled them to find the solutions with lesser effort and time when compared to the rest of the groups. The possible reason for this could be that the proposed system had gathered sufficient information from the completion of the previous sessions of the earlier three groups, thereby, enabling the system to produce a set of the better recommendation of links for Group 4 to find their solutions.

On the contrary, when Group 1 was using the proposed system, it had not accumulated enough browsing information for the system to produce recommendations. The system had only provided the students with default search results returned by the Google search engine. Group 1 had to search for answers within this default search results. Therefore, the level of satisfaction in terms of perceived ease of usefulness and ease of use when using the proposed system was exhibited modestly by Group 1. Similarly, the higher level of satisfaction was observed in Group 3 as compared to Group 2. Nevertheless, most students had exhibited a high level of perception on the ease of use and usefulness while using the proposed Personalised Web Search system for e-Learning. Figure 6.4 shows that most of the students identified the proposed system to be very easy to use and useful. It is also evident that the group-based recommendation approach used in our proposed system enhanced and so improved the students' learning performance. The proposed system had helped them to find the relevant personalised learning materials with ease; therefore, they were able to solve the given problems successfully and effortlessly.

The ANOVA results generated by the TAM questionnaire for all three levels (Beginner, Intermediate and Master) of students' are shown in Table 6.4. The result showed that the mean ratings of the perceived "ease of use" ($F = 0.0264$, $p<1$) for Beginner, Intermediate and Master level profiles are 4.03, 4.02 and 4.07 respectively. The mean ratings for "perceived usefulness" are 4.11, 4.01 and 4.02 for each level of students individually. The results indicate that there was no significant difference between the means of the three levels of students. Students from all the levels expressed a similar degree of satisfaction on an average. Particularly the novice learners revealed a slightly higher degree of satisfaction compared with other levels of students. Figure 6.5 demonstrates the most of the students showed satisfaction while using the proposed system regardless of their learning profile.

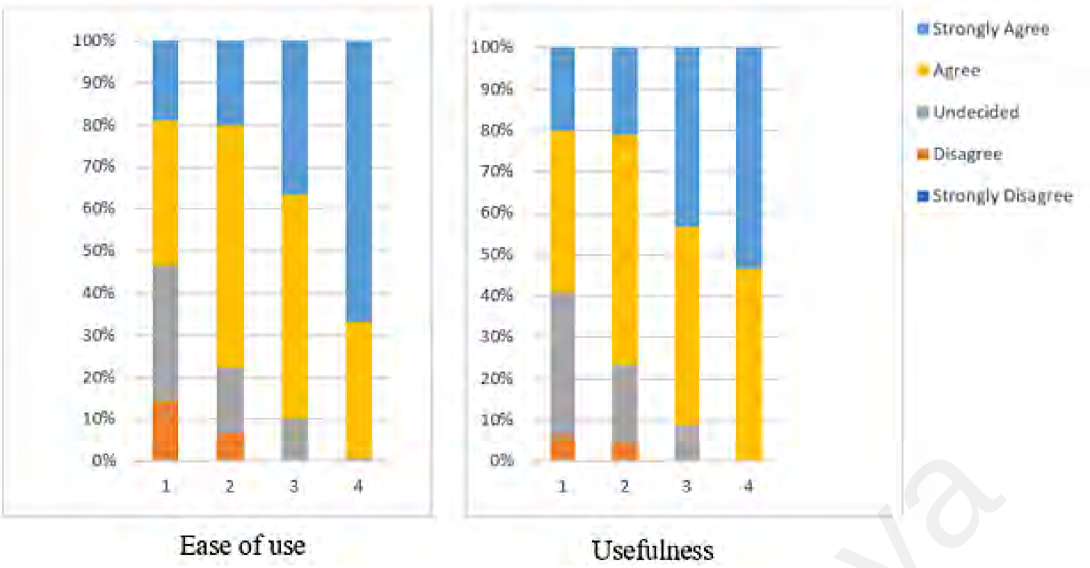

Figure 6.4: The overall result of the students' perception towards the proposed system (experimental groups)

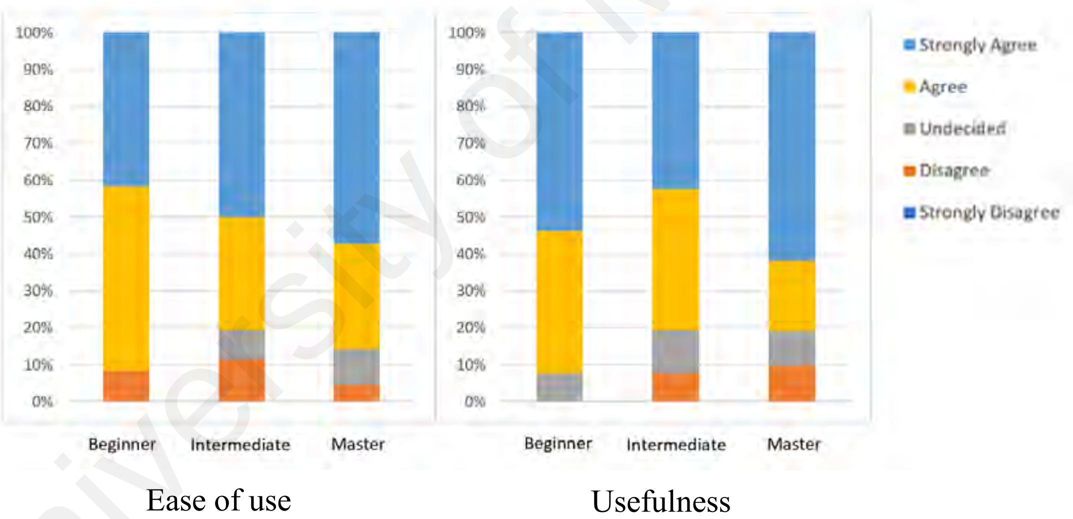

Figure 6.5: The overall result of the students' perception towards the proposed system (three profile levels)

6.6.3 Analysis of Searching Time

Results obtained from the experiment showed the effectiveness of our proposed system in terms of searching time. This implies how quickly students were able to find their learning materials from the search engine. The analysis was generated to scrutinise the period of time the students from each group took to search for answer materials in order

Table 6.5: ANOVA results of the four experimental groups (searching time)

Scale	Groups	N	Mean	SD	F	Post hoc tests	CI
Search Time	Group 1 (a)	15	34.60	3.97	25.89	a > b	95
	Group 2 (b)	15	30.73	5.53		b > c	
	Group 3 (c)	15	24.40	4.50		c > d	
	Group 4 (d)	15	21.26	4.21			

*p<0.05

Table 6.6: ANOVA results of the three profile levels (searching time)

Scale	Groups	N	Mean	SD	F	Post hoc tests	CI
Search Time	Beginner (a)	13	33.44	6.69	6.15	a > b	95
	Intermediate (b)	26	31.49	5.76		b > c	
	Master (c)	21	26.78	5.49		c > d	

*p<0.05

to complete their given tasks. Table 6.5 shows the ANOVA results of the searching time for the four groups. The mean and standard deviation of the searching time for all the four groups were 34.60 and 3.97, 30.73 and 5.53, 24.40 and 4.50 and 21.26 and 4.21, respectively. It is observed that there was a decrease in mean searching time ranging from Group 1 to Group 4. For example, the mean searching time for Group 1 was 3.87 minutes longer than Group 2, and the mean searching time for Group 3 was 6.33 minutes shorter than Group two, and so on. Evidently, Group 4 got their desired materials using our proposed system more expeditiously than the remaining groups. The possible reason causing this can be attributed to the fact that each group received better recommendations over time, for finding worthy materials to solve the given problems. This is due to the higher quality recommendations generated by the system when it had received sufficient searching history from the previous groups, hence by the time Group 4 was using the prototype, adequate history had been received by the system to generate more useful and personalised links. As explained above, Group 1 was deprived of the benefits due to a lack of gathered information to facilitate the recommendation process. Consequently, they had to depend on the default links returned by the Google search engine. Only a few participants in Group 1 had benefited from the run-time searching history provided by other participants within the same group in the same session. In comparison, Group 4 utilised the shortest time to solve the problems as they were able to receive quality recommendations from the proposed system. Furthermore, the ANOVA results of the searching time for three levels of students are shown in Table 6.6. The average searching time for each level of students: Beginner, Intermediate and Master are 33.44, 31.49 and 26.78 seconds respectively. As expected, students from Beginner levels took the highest time to complete the tasks while students from Master level took the lowest. There is a significant difference in searching time among each level of students. However, the overall results revealed that the proposed Personalised Web search system for e-Learning using the group-based recommendation approach had a significant impact in reducing students' search time for e-Learning materials through the Web search engine regardless of their profile.

6.6.4 Analysis of Searching Performance

In the experiment, the participants' link selection was also monitored so as to obtain information about the participants' preferences towards the links displayed in each search results. Figure 6.6 shows the total number of clicks noted on the links returned by the Google search engine and the recommended links of each experimental group. It was noticed that the number of clicks identified on the recommended links increases while the number of clicks on Google returned links decreases when Group 1 is compared to Group 4. This outcome reveals an inverse relationship between recommendation links and the Google returned links. One possible reason for the high number of clicks on the links displayed by Google is due to the students' familiarity with the Google default returned search results when compared to the recommended links. The noteworthy finding obtained from this experiment indicates that there is an incremental interest to select the recommended links. As mentioned in the earlier part of this study, as the participation of each group increases, their browsing histories also become richer, thereby making it possible for the proposed system to recommend links that were personalised to their profiles. Furthermore, we also compared the performance of the students from three profile levels. The investigated results (Figure 6.7) revealed that the rate of the recommendation links clicked (total number of recommendation links clicked/total number of links clicked including Google and recommendation links) from the students of Beginner level is the highest which is around 29.69% followed by Intermediate 21.57% and Master 17.9%. This is a significant indication that the beginner level of students benefited most in the provided personalised recommended links to find their desired solutions.

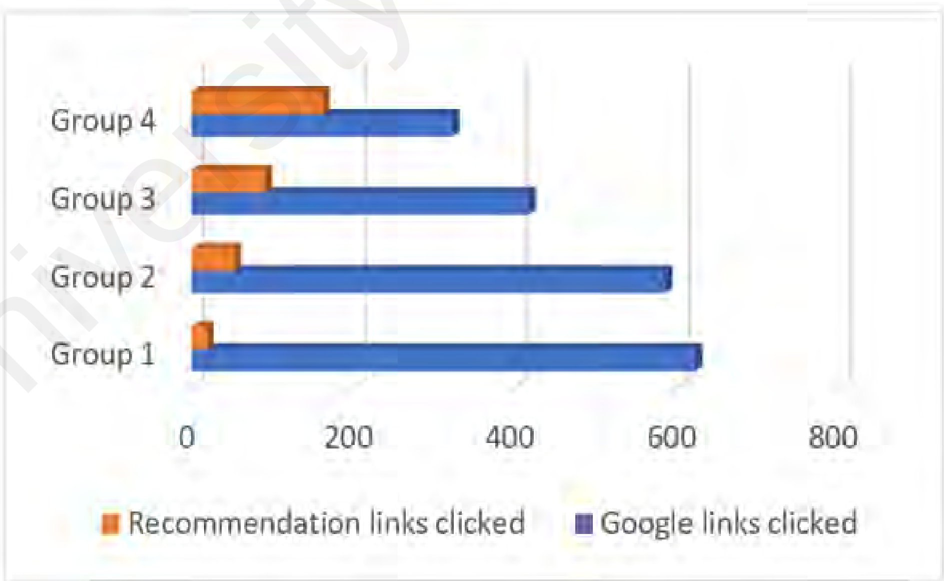

Figure 6.6: Links clicked by the participants (four experimental groups)

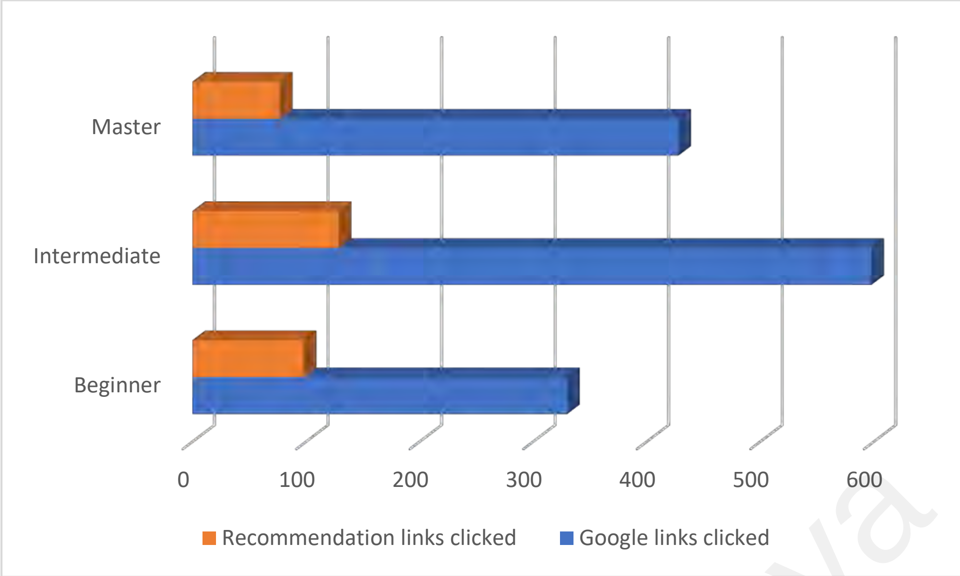

Figure 6.7: Links clicked by the participants (three profile levels)

6.7 Chapter Summary

This chapter has explained the outcome generated from the experiment conducted on the four groups of participants and the results drawn from the survey administered. The results clearly indicated that the proposed personalised Web search system contributes to e-learning by ensuring that the recommended links provided by our proposed system enable users to have more benefits. The next chapter concludes this study.

CHAPTER 7: CONCLUSION

The last chapter of this dissertation summarises the entire study and the findings generated. This chapter also presents a discussion about the limitation of this study besides making recommendations for future work.

7.1 Research Summary

The World Wide Web (WWW) acts as a large database repository from which users are able to gain the information which they are searching for. A Web search engine is a specialised software program that is used to help users to mine particular information from billions of databases or open directories. Presently, students use various Web search engines extensively to search for the relevant e-Learning materials. Unfortunately, the current Web search engines tend to deliver the same learning materials to all these students regardless of their learning differences. Undoubtedly, the same set of learning materials may not be suitable for every student since every student has various levels of absorption and understanding. Web search results that are personalised based on students' learning profile are, therefore, more suitable and more necessary.

In this study, a novel and Personalised Web search approach for e-Learning was proposed so that it will present personalised recommendations of search results to e-Learners based on their learning profiles and their group relationship with others. To the best of our knowledge, the current study and its aim would be the first attempt to locate and record students' academic records and their learning behaviours within a system so that the dynamic students' profile can be used by the system to deliver Personalised Web search recommendations. The T-Score formula was used to standardise the raw scores of the students' academic records. A knowledge point (KP) was then calculated by averaging the standard academic T-Scores for each student. The students' learning behavioural information was then extracted from their Web browsing history and session logs while they were using the proposed system. Their behavioural information was then used to classify their level of interest into four ranks: Low, Low Medium, Medium and High by using fuzzification rules. Finally, the students' knowledge point and interest level were utilised to model the dynamic students' profile by using the C4.5 algorithm. Each of the students belongs to any of the three classes: Beginner, Intermediate and Master.

In order to re-rank the Web search results in a personalised way that meets the individual student's needs, the Groupization technique was applied in this study. The technique uncovered the students with shared interest within similar groups by harvesting their profiles to improve the personalisation of the Web search results (in document ranking level). However, this shared information regarding students' group relationship is usually insignificant or unavailable to the public Web search engines. Consequently, they are unable to obtain information regarding the relationships among the group members due to the unavailability of public data and privacy concern. In contrast, the proposed Personalised Web search approach that uses the group-based recommendation approach can identify different user profiles within the homogeneous group of users (students) within an educational institution environment. This can strengthen the approach in delivering more accurate personalised recommendations derived from the search results within the e-Learning environment.

To measure the acceptability and reliability of the proposed system, an experiment was conducted with a group of 60 second year undergraduate students. Four groups were created by randomly selecting 15 students regardless of their academic records and they were then placed into four groups and they were monitored in four different sessions. The users' level of satisfaction in terms of ease of use and usefulness of the proposed system was measured. In addition, the searching performance and time are taken by the users of the four groups while using the proposed system were also evaluated. The results revealed that about 98% (66% strongly agree and 32% agree) of the students from Group 4 found the system easy to use followed by 88.5% from Group 3, 78% from Group 2 and 54% Group 1 respectively. Similarly, most of the students from all the four groups found the system to be very useful for searching the learning materials particularly the Beginner level of students. Moreover, to measure the level of acceptance among the different levels of students such as Beginner, Intermediate and Master, it was found that more than 80% students from each level found the system easy to use and its usefulness while searching for learning materials. It is noted that especially Beginner levels of learners revealed the highest level of satisfaction. Further, it was observed that the participants' searching time for finding learning materials significantly decreased and their searching performance had improved when observed based on groups. It significantly improved the performance of the Master level of learners by reducing searching time. This implies that the proposed Personalised Web search approach using the group-based recommendation approach was able to make the searching process for learning materials more effective and more accessible, thereby less exhaustive regardless of the level of student's profile. In other words, it offered better support and it also helped to promote a successful collaborative Web-based learning attitude.

7.2 Overview of the Research Findings

The aim of this study was to provide students with a recommendation list of Web search results that are personalised according to their individual learning profiles. The work developed in this study was motivated by the fact that using the same search keyword will only result in the current Web search engine returning the same list of search results for all the students regardless of their differences. Thus, three main objectives were set for this study as specified in Section 1.4. This section discusses the findings of each objective.

7.2.1 To investigate the strengths and weaknesses of the current Web search in accommodating the needs of the individual students for e-learning materials

For this objective, literature related to Web search in e-Learning, existing research work in specialised search, personalisation and recommendations noted in current literature were studied and analysed. The analysis revealed that traditional Web search engines are very much used by students who use the Web search engines extensively as an e-Learning tool, besides the traditional classroom learning environment. However, these systems provide very limited support which offers Personalised Web search results to them. Furthermore, most previous research works were conducted to study how to deliver personalised learning materials in various e-Learning applications. Only a few studies were found to discuss other popular Web search engines as an e-Learning tool. We also note that thus far, no research has been conducted in offering personalised search recommendations to students who use the most popular Web search engines for their e-learning purposes. Augmenting the students' dynamic profiling and investigating the group relationships among users (students) can leverage the personalised recommendation process in Web search.

7.2.2 To investigate the strengths and weaknesses of the current Web search in accommodating the needs of the individual students for e-learning materials

To achieve this objective, we proposed a framework that can act as a middleware between the Web search engine and the institution's e-Learning portal for delivering Personalized Web search results. The framework contains two important features which are dynamic student profiling and the group-based content selection and re-ranking. The dynamic students' profiling technique was based on their academic records and learning behaviours. This can enable the system to appraise the individual learning needs of the students and to classify the students into three classes (i.e. Beginner, Intermediate and Master). The Groupization technique that we deployed was able to re-rank the Web search results to match the individual student's profile. The developed prototype that was reported in Chapter 5 can deliver Personalised Web search results for students by using the group-based recommendation approach.

7.2.3 To evaluate the feasibility of the proposed Web search approach for e-Learning

To achieve this objective, we conducted four consecutive experiments on four different groups of undergraduate students (Groups 1-4). Each group comprised of 15 students. The results of the experiment reported in Chapter 6 indicate that students from Group 4 benefitted the most from the personalised recommendation links presented by the proposed system. They gained better learning comprehension than the earlier three groups through the Personalised Web search results. Notably, students from the preceding three groups demonstrated better learning performance as compared to their immediate preceding groups (i.e. Group 3 achieved better performance than Group 2, and so on). This implies that the group-based recommender system is helpful in improving the effectiveness of the returned results of the students' Web search. By recommending a list of search results that are personalised to the student's learning profile, the system helped to accelerate the process of finding learning materials that matched the needs of each individual student, hence the learning problem could be solved faster.

The survey results generated from the questionnaire that was adapted from TAM revealed that the highest percentage or 52% of the students from Group 4 had strongly agreed and 48% had agreed that the proposed system is useful to them. This is followed by Group 3 (92%), Group 2 (76%) and Group 1 (59%) (refer Figure 6.4). In terms of its ease of use, 98% of the students from Group 4 were satisfied with the proposed system (66% strongly agree and 32% agree). This is followed by Group 3, 2 and 1 with 88.5%, 78% and 54%, respectively (refer Figure 6.4). Moreover, around 80% of students from each level found the system easy to use and its usefulness. These results suggest that the personalised recommender system using the group-based recommendation approach in Web search is able to ease the process of searching for e-Learning materials regardless of learning level.

7.3 Research Contribution

This research contributes to the e-Learning community, particularly for those involved with Web search, especially when popular Web search engines are being used to search for e-Learning materials. The findings of this study provide a novel approach of the Personalised Web search for e-Learning by using the group-based recommendation approach. This outcome augments the Web search engines that can offer better support for e-learning and it can also promote the successful collaborative Web-based learning attitude. It also adds values in modelling the dynamic students' profiling by using their academic records and learning behaviours. The group-based recommendation approach also leverages the contents' re-ranking mechanism in order to present personalised recommendations that match the individual student's profiles.

7.4 Limitation

In this study, the Google API was utilised to integrate the Google search engine with our developed prototype. A point to note here is that Google API imposes a number of limitations such as the 100 free search queries per day, differences in the order of search results, and differences in presentations. Nevertheless, these limitations can be overcome by requesting support from Google for educational purposes. The most challenging task, however, is to encourage students to use the Web search engine through an institutional e-learning portal.

7.5 Future Works

The proposed novel Web search technique for e-Learning using the group-based recommendation approach offers students a greater learning comprehension when used with the personalised Web search experience. In order to model a rich profile of the student, it is necessary to integrate the student's social media data with the profiling technique so that it helps to improve the personalised recommendation accuracy. This should be included in future works. In addition, it would be interesting to further explore how e-Learning extensions or add-ons can be incorporated into browsers for the purpose of providing personalised recommended links to e-Learners.

REFERENCES

6th Annual LMS Data Update | edutechnica. (2018). Retrieved April 4, 2019, from http://edutechnica.com/2018/10/06/6th-annual-lms-data-update/

Alharthi, A. D., Spichkova, M., & Hamilton, M. (2015). Requirements Engineering Aspects of ELearning Systems. In *Proceedings of the ASWEC 2015 24th Australasian Software Engineering Conference. ACM* (pp. 132–133). https://doi.org/10.1145/2811681.2817756

Amer-Yahia, S., Roy, S., & Chawlat, A. (2009). Group recommendation: Semantics and efficiency. *Proceedings of the VLDB Endowment*, *2*(1), 754–765. https://doi.org/10.1.1.151.4805

Amershi, S., & Morris, M. R. (2008). CoSearch. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08* (p. 1647). New York, New York, USA: ACM Press. https://doi.org/10.1145/1357054.1357311

Anagnostopoulos, A., Atassi, R., Becchetti, L., Fazzone, A., & Silvestri, F. (2017). Tour recommendation for groups. *Data Mining and Knowledge Discovery*, *31*(5), 1157–1188. https://doi.org/10.1007/s10618-016-0477-7

Anuradha, C. (2012). An effective method for personalized image search using query expansion. *Middle - East Journal of Scientific Research*, *12*(12), 1741–1744. https://doi.org/10.5829/idosi.mejsr.2012.12.12.1215

Baguley, M., Danaher, P. A., Davies, A., De George-Walker, L., Jones, J. K., Matthews, K. J., … Arden, C. H. (2014). *Educational Learning and Development*. London: Palgrave Macmillan UK. https://doi.org/10.1057/9781137392848

Ben Ahmed, E., Nabli, A., & Gargouri, F. (2012). *Performing groupization in data warehouses: which discriminating criterion to select? Natural Language Processing and Information Systems* (Vol. 7337). Springer.

Ben Ahmed, E., Nabli, A., & Gargouri, F. (2013). A new semi-supervised hierarchical active clustering based on ranking constraints for analysts groupization. *Applied Intelligence*, *39*(2), 236–250. https://doi.org/10.1007/s10489-012-0407-3

Brown, R. (2010). Collaborative learning. In *Collaborative learning BT - Collaborative learning*.

Campbell, I. D., Lindskog, S., & White, A. I. (1975). A study of the histidine residues of human carbonic anhydrase C using 270 MHz proton magnetic resonance.

Journal of Molecular Biology, *98*(3), 597–614. https://doi.org/10.5171/2013.

Capra, R., Arguello, J., Crescenzi, A., & Vardell, E. (2015). Differences in the Use of Search Assistance for Tasks of Varying Complexity. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '15* (pp. 23–32). https://doi.org/10.1145/2766462.2767741

Capra, R., Marchionini, G., Oh, J. S., Stutzman, F., & Zhang, Y. (2007). Effects of structure and interaction style on distinct search tasks. In *Proceedings of the 2007 conference on Digital libraries - JCDL '07* (p. 442). https://doi.org/10.1145/1255175.1255267

Carey, J. J., & Delaney, M. F. (2010). T-scores and Z-scores. *Clinical Reviews in Bone and Mineral Metabolism*. https://doi.org/10.1007/s12018-009-9064-4

Cecchino, N. J. (2010). Google Scholar. *Journal of the Medical Library Association : JMLA*, *98*(4), 320–321. https://doi.org/10.3163/1536-5050.98.4.016

Chen, C. M., Lee, H. M., & Chen, Y. H. (2005). Personalized e-learning system using Item Response Theory. *Computers and Education*, *44*(3), 237–255. https://doi.org/10.1016/j.compedu.2004.01.006

Clark, R. C., & Mayer, R. E. (2011). *e-Learning and the Science of Instruction*. *e-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning: Third Edition*. San Francisco, CA, USA: Pfeiffer. https://doi.org/10.1002/9781118255971

Cuevas, A., Febrero, M., & Fraiman, R. (2004). An anova test for functional data. *Computational Statistics and Data Analysis*, *47*(1), 111–122. https://doi.org/10.1016/j.csda.2003.10.021

Curlango-Rosas, C., Ponce, G. a., & Lopez-Morteo, G. a. (2011). A Specialized Search Assistant for Learning Objects. *ACM Transactions on the Web*, *5*(4), 1–29. https://doi.org/10.1145/2019643.2019648

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Dogruer, N., Eyyam, R., & Menevis, I. (2011). The use of the internet for educational purposes. *Procedia - Social and Behavioral Sciences*, *28*, 606–611. https://doi.org/10.1016/j.sbspro.2011.11.115

Faulkner, K. G., Von Stetten, E., & Miller, P. (1999). Discordance in patient classification using T-scores. *Journal of Clinical Densitometry*, *2*(3), 343–350. https://doi.org/10.1385/JCD:2:3:343

Felder, R. M., & Brent, R. (2005). Understanding student differences. In *Journal of Engineering Education* (Vol. 94, pp. 57–72). https://doi.org/10.1002/j.2168-9830.2005.tb00829.x

Fischer, G. (2001). User modeling in human-computer interaction. *User Modeling and User-Adapted Interaction*, *11*(1–2), 65–68. https://doi.org/10.1023/A:1011145532042

Gavalas, D., Konstantopoulos, C., Mastakas, K., & Pantziou, G. (2014). Mobile recommender systems in tourism. *Journal of Network and Computer Applications*, *39*, 319–333. https://doi.org/10.1016/j.jnca.2013.04.006

Ge, S., Dou, Z., Jiang, Z., Nie, J.-Y., & Wen, J.-R. (2018). Personalizing Search Results Using Hierarchical RNN with Query-aware Attention. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management - CIKM '18* (pp. 347–356). New York, New York, USA: ACM Press. https://doi.org/10.1145/3269206.3271728

Giordani, B., Novak, B., Sikorskii, A., Bangirana, P., Nakasujja, N., Winn, B. M., & Boivin, M. J. (2015). Designing and evaluating Brain Powered Games for cognitive training and rehabilitation in at-risk African children. *Global Mental Health*, *2*(May), e6. https://doi.org/10.1017/gmh.2015.5

Guha, R., Gupta, V., Raghunathan, V., & Srikant, R. (2015). User Modeling for a Personal Assistant. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining - WSDM '15* (pp. 275–284). New York, New York, USA: ACM Press. https://doi.org/10.1145/2684822.2685309

Hassan, A., Jones, R., & Klinkner, K. L. (2010). Beyond DCG. In *Proceedings of the third ACM international conference on Web search and data mining - WSDM '10* (p. 221). New York, New York, USA: ACM Press. https://doi.org/10.1145/1718487.1718515

Hassan, S., & Mihalcea, R. (2011). Learning to identify educational materials. *ACM Transactions on Speech and Language Processing*, *8*(2), 1–18. https://doi.org/10.1145/2050100.2050101

Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, *22*(1), 5–53. https://doi.org/10.1145/963770.963772

Hssina, B., Merbouha, A., Ezzikouri, H., & Erritali, M. (2014). A comparative study of decision tree ID3 and C4.5. *International Journal of Advanced Computer Science and Applications*, *4*(2). https://doi.org/10.14569/SpecialIssue.2014.040203

Huang, C. K., Chien, L. F., & Oyang, Y. J. (2003). Relevant term suggestion in interactive web search based on contextual information in query session logs. *Journal of the American Society for Information Science and Technology*, *54*(7), 638–649. https://doi.org/10.1002/asi.10256

Ingleby, E. (2012). Research methods in education. *Professional Development in Education*, *38*(3), 507–509. https://doi.org/10.1080/19415257.2011.643130

Introduction, a S. (1994). Neural networks referees in 1993. *Neural Networks*, *7*(1), xiii–xiv. https://doi.org/10.1016/0893-6080(94)90051-5

iSEEK - Education. (2009.). Retrieved December 12, 2017, from http://education.iseek.com/iseek/home.page

Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, *16*(3), 261–273. https://doi.org/10.1016/j.eij.2015.06.005

Islam, M. J., Wu, Q. M. J., Ahmadi, M., & Sid-Ahmed, M. A. (2007). Investigating the Performance of Naive- Bayes Classifiers and K- Nearest Neighbor Classifiers. In *2007 International Conference on Convergence Information Technology (ICCIT 2007)* (pp. 1541–1546). IEEE. https://doi.org/10.1109/ICCIT.2007.4420473

Iverson, B. (2017). Detailed Grading Policy. Retrieved November 30, 2017, from http://iverson.cm.utexas.edu/courses/310N/MainPagesSp06/GradingPolicy.html

Iverson, G. L. (2011). T Scores. In *Encyclopedia of Clinical Neuropsychology* (pp. 2459–2460). New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-79948-3_1254

Kacem, A., Boughanem, M., & Faiz, R. (2017). Emphasizing temporal-based user profile modeling in the context of session search. In *Proceedings of the Symposium on Applied Computing - SAC '17* (pp. 925–930). New York, New York, USA: ACM Press. https://doi.org/10.1145/3019612.3019693

Keenoy, K., & Levene, M. (2005). Personalisation of Web Search. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 201–228). https://doi.org/10.1007/11577935_11

Kelly, D., & Azzopardi, L. (2015). How many results per page? In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '15* (pp. 183–192). New York, New York, USA: ACM Press. https://doi.org/10.1145/2766462.2767732

Klašnja-Milićević, A., Vesin, B., Ivanović, M., Budimac, Z., & Jain, L. C. (2017). *E-Learning Systems* (Vol. 112). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-41163-7

Krus, D. J., & Krus, P. H. (1977). Lost: McCall's T Scores: Why? *Educational and Psychological Measurement*, *37*(1), 257–261. https://doi.org/10.1177/001316447703700134

Kumar, A., & Ashraf, M. (2015). Personalized web search engine using dynamic user profile and clustering techniques. In *Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on* (pp. 2105–2108).

Lakshmi, B. N., Indumathi, T. S., & Ravi, N. (2016). A Study on C.5 Decision Tree Classification Algorithm for Risk Predictions During Pregnancy. *Procedia Technology*, *24*, 1542–1549. https://doi.org/10.1016/j.protcy.2016.05.128

Leung, K. W.-T., Lee, D. L., Ng, W., & Fung, H. Y. (2012). A framework for personalizing web search with concept-based user profiles. *ACM Transactions on Internet Technology*, *11*(4), 1–29. https://doi.org/10.1145/2109211.2109214

Leung, K. W. T., Lee, D. L., & Lee, W. C. (2010). Personalized web search with location preferences. In *Proceedings - International Conference on Data Engineering* (pp. 701–712). https://doi.org/10.1109/ICDE.2010.5447911

Li, Y.-J., Li, Q.-S., & Lin, Y.-S. (2017). A Personalized Result Recommendation Method based on Communities. In *Proceedings of the 2017 International Conference on Data Mining, Communications and Information Technology - DMCIT '17* (pp. 1–5). New York, New York, USA: ACM Press. https://doi.org/10.1145/3089871.3089885

Li, Y., Luo, J., & Mei, T. (2014). Personalized image recommendation for web search engine users. In *Proceedings - IEEE International Conference on Multimedia and Expo* (Vol. 2014-Septe). https://doi.org/10.1109/ICME.2014.6890327

Liu, B. (2006). *Personal evaluations of search engines: Google, Yahoo! and MSN. Department of Computer Science University of Illinois at Chicago*. Retrieved from https://www.cs.uic.edu/~liub/searchEval/SearchEngineEvaluation.htm

Lts, C. (2000). *IEEE Learning Technology Standards Committee (LTSC): IEEE P1484.*

12 Learning Object Metadata Working Group (home page).

Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: A survey. *Decision Support Systems*, *74*, 12–32. https://doi.org/10.1016/j.dss.2015.03.008

Makvana, K., Jay, P., Shah, P., & Thakkar, A. (2016). An Approach to identify semantic relations between user's queries in text retrieval. In *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies - ICTCS '16* (pp. 1–6). New York, New York, USA: ACM Press. https://doi.org/10.1145/2905055.2905271

Makvana, K., Shah, P., & Shah, P. (2014). A novel approach to personalize web search through user profiling and query reformulation. In *2014 International Conference on Data Mining and Intelligent Computing, ICDMIC 2014*. https://doi.org/10.1109/ICDMIC.2014.6954221

Marcato, E., & Scala, E. (2012). Moodle. In *Handbook of Research on Didactic Strategies and Technologies for Education* (pp. 107–116). IGI Global. https://doi.org/10.4018/978-1-4666-2122-0.ch010

Marin, M., Gil-Costa, V., Bonacic, C., & Inostrosa, A. (2017). Simulating Search Engines. *Computing in Science & Engineering*, *19*(1), 62–73. https://doi.org/10.1109/MCSE.2017.8

Microsoft. (2013). Bing in the Classroom - Microsoft Education. Retrieved November 30, 2017, from https://www.bing.com/classroom

Morris, M. G., & Dillon, A. (1997). The Influence of User Perceptions on Software Utilization: Application and Evaluation of a Theoretical Model of Technology Acceptance. *IEEE Software*, *14*(4), 58–6. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.12.9088&rep=rep1&type=pdf

Morris, Meredith R, Teevan, J., & Bush, S. (2008). Enhancing collaborative web search with personalization: groupization, smart splitting, and group hit-highlighting. In *Computer Supported Cooperative Work* (p. 3). https://doi.org/10.1145/1460563.1460640

Morris, Meredith Ringel. (2008). A Survey of Collaborative Web Search Practices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)* (pp. 1657--1660). https://doi.org/10.1145/1357054.1357312

Moss, K. (2012). The Entity-Relationship model. In *IEEE Global Engineering Education Conference, EDUCON*.

https://doi.org/10.1109/EDUCON.2012.6201182

Oracle. (2016). MySQL | The Most Popular Open-Source Database | Oracle.

Park, M. H., Park, H. S., & Cho, S. B. (2008). Restaurant recommendation for group of people in mobile environments using probabilistic multi-criteria decision making. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-540-70585-7_13

Pignatiello, M. F., Camp, C. J., & Rasar, L. A. (1986). Musical Mood Induction. An Alternative to the Velten Technique. *Journal of Abnormal Psychology*, *95*(3), 295–297. https://doi.org/10.1037/0021-843X.95.3.295

Premlatha, K. R., Dharani, B., & Geetha, T. V. (2016). Dynamic learner profiling and automatic learner classification for adaptive e-learning environment. *Interactive Learning Environments*, *24*(6), 1054–1075. https://doi.org/10.1080/10494820.2014.948459

Priyanka C. Ghegade, P. V. W. (2014). A Survey of Personalized Web Search in Current Techniques. *International Journal of Computer Science and Information Technologies*, *5*(6), 7945–7947.

Qiu, F., & Cho, J. (2006). Automatic identification of user interest for personalized search. In *Proceedings of the 15th international conference on World Wide Web* (pp. 727–736). https://doi.org/10.1145/1135777.1135883

Quijano-Sánchezz, L., Recio-García, J. A., & Díaz-Agudo, B. (2011). Group recommendation methods for social network environments. In *Recommender Systems and the Social Web* (p. 24).

Quinlan, J. R. (1992). *C4.5: Programs for Machine Learning*. *Morgan Kaufmann San Mateo California* (Vol. 1). https://doi.org/10.1016/S0019-9958(62)90649-6

Rahman, M. M., Abdullah, N. A., & Aurangozeb, F. (2017). A framework for designing a personalised web-based search assistant tool for eLearning. In *2017 5th International Conference on Information and Communication Technology, ICoIC7 2017*. https://doi.org/10.1109/ICoICT.2017.8074712

Rizk, N. M., Gheith, M. H., & Nasr, E. S. (2016). Requirements' elicitation needs for eLearning Systems. In *2016 12th International Computer Engineering Conference (ICENCO)* (pp. 142–147). IEEE. https://doi.org/10.1109/ICENCO.2016.7856459

Rokach, L., & Maimon, O. (2010). Decision Trees. In *Data Mining and Knowledge*

Discovery Handbook (pp. 165–192). New York: Springer-Verlag. https://doi.org/10.1007/0-387-25465-X_9

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning*, *4*(1), 69–92. https://doi.org/10.1007/s11412-008-9054-4

S, S., K, S., & G.S, M. (2014). Application of Fuzzy Logic for User Classification in Personalized Web Search. *International Journal on Cybernetics & Informatics*, *3*(3), 23–49. https://doi.org/10.5121/ijci.2014.3303

Said, A., Berkovsky, S., & De Luca, E. W. (2011). Group recommendation in context. In *Proceedings of the 2nd Challenge on Context-Aware Movie Recommendation - CAMRa '11* (pp. 2–4). New York, New York, USA: ACM Press. https://doi.org/10.1145/2096112.2096113

Santos, J. R. A. (1999). Cronbach's alpha: A tool for assessing the reliability of scales. *Journal of Extension*, *37*(2), 1–5.

Sarker, A., Ginn, R., Nikfarjam, A., O'Connor, K., Smith, K., Jayaraman, S., … Gonzalez, G. (2015). Utilizing social media data for pharmacovigilance: A review. *Journal of Biomedical Informatics*, *54*, 202–212. https://doi.org/10.1016/j.jbi.2015.02.004

Sathiyabama, T., & Vivekanandan, K. (2011). Personalized Web Search Techniques - A Review. *Global Journal of Computer Science and Technology*, *11*(12).

Shafiq, O., Alhajj, R., & Rokne, J. G. (2010). Community Aware Personalized Web Search. In *2010 International Conference on Advances in Social Networks Analysis and Mining* (pp. 351–355). IEEE. https://doi.org/10.1109/ASONAM.2010.85

Sherman, C. (2001). Consumer Watchdog Files Complaint Against Eight Search Engines for "Crass Commercialism." Retrieved November 30, 2017, from http://newsbreaks.infotoday.com/NewsBreaks/Consumer-Watchdog-Files-Complaint-AgainstEight-Search-Engines-for-Crass-Commercialism-17544.asp

Stern, D. H., Herbrich, R., & Graepel, T. (2009). Matchbox. In *Proceedings of the 18th international conference on World wide web - WWW '09* (p. 111). New York, New York, USA: ACM Press. https://doi.org/10.1145/1526709.1526725

Suguna, S., Sundaravadivelu, V., & Gomathi, B. (2016). A novel semantic approach in E-learning information retrieval system. In *Proceedings of 2nd IEEE International Conference on Engineering and Technology, ICETECH 2016*. https://doi.org/10.1109/ICETECH.2016.7569374

Summary, E. (2013). Executive Summary. *New Directions for Youth Development*, *2013*(140), 5–8. https://doi.org/10.1002/yd.20075

Tamine, L., & Soulier Pierre, L. (2016). *Collaborative Information Retrieval: Concepts, Models and Evaluation*. Springer. Retrieved from http://ecir2016.dei.unipd.it/presentations/CIR_Tutorial.pdf

Tatroe, K., MacIntyre, P., & Lerdorf, R. (2013). *Programming PHP - Creating Dynamic Web Pages*. *O'Reilly*.

Teevan, J., Dumais, S. T., & Horvitz, E. (2005). Personalizing search via automated analysis of interests and activities. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '05* (p. 449). https://doi.org/10.1145/1076034.1076111

Teevan, J., Morris, M. R., & Bush, S. (2009). Discovering and using groups to improve personalized search. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining - WSDM '09* (p. 15). New York, New York, USA: ACM Press. https://doi.org/10.1145/1498759.1498786

Upton, D. (2007). *Codeigniter for Rapid Php Application Development*. *Image Rochester NY*.

Wang, H., & Wong, K. (2013). Recommendation-Assisted Personal Web. In *2013 IEEE Ninth World Congress on Services* (pp. 136–140). IEEE. https://doi.org/10.1109/SERVICES.2013.20

Watkins, C. (2009). Collaborative learning. *Learning*, *1*, 37–54. Retrieved from http://eprints.ioe.ac.uk/2818/

Wei, X., & Yan, J. (2009). Learner Profile Design for Personalized E-Learning Systems. In *2009 International Conference on Computational Intelligence and Software Engineering* (pp. 1–4). IEEE. https://doi.org/10.1109/CISE.2009.5363560

White, R. W., Chu, W., Hassan, A., He, X., Song, Y., & Wang, H. (2013). Enhancing personalized search by mining and modeling task behavior. In *In Proceedings of the 22nd international conference on World Wide Web (WWW '13)* (pp. 1411–1420). https://doi.org/10.1145/2488388.2488511

Witten, I. H., Frank, E., & Hall, M. a. (2001). No Title. *Angewandte Chemie International Edition*, *40*(6), 9823. https://doi.org/10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C

Yathongchai, C., Angskun, T., Yathongchai, W., & Angskun, J. (2013). Learner classification based on learning behavior and performance. In *2013 IEEE Conference on Open Systems (ICOS)* (pp. 66–70). IEEE. https://doi.org/10.1109/ICOS.2013.6735050

Zamani, H., Bendersky, M., Wang, X., & Zhang, M. (2017). Situational Context for Ranking in Personal Search. In *Proceedings of the 26th International Conference on World Wide Web - WWW '17* (pp. 1531–1540). New York, New York, USA: ACM Press. https://doi.org/10.1145/3038912.3052648

Zemede, B. A., & Gao, B. J. (2017). Personalized search with editable profiles. In *2017 IEEE International Conference on Big Data (Big Data)* (pp. 4872–4874). Boston, MA, USA: IEEE. https://doi.org/10.1109/BigData.2017.8258572

Zenetti, G., Bijmolt, T. H. A., Leeflang, P. S. H., & Klapper, D. (2014). Search Engine Advertising Effectiveness in a Multimedia Campaign. *International Journal of Electronic Commerce*, *18*(3), 7–38. https://doi.org/10.2753/JEC1086-4415180301

Zweihorn, Z. J. (2006). Searching for confusion: The initial interest confusion doctrine and its misapplication to search engine sponsored links. *Cornell Law Review*, *91*(6), 1343–1381.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Conference Proceeding

M. M. Rahman, N. A. Abdullah and F. Aurangozeb, "A framework for designing a personalised web-based search assistant tool for eLearning," 2017 5th International Conference on Information and Communication Technology (ICoIC7), Melaka, 2017, pp. 1-6.

Journal

Rahman, M. M., & Abdullah, N. A. (2018). A Personalized Group-Based Recommendation Approach for Web Search in E-Learning. IEEE Access, 6, 34166-34178.