

ADAPTIVE LEARNING METHOD FOR WEB RECOMMENDATION

A PROJECT REPORT

Submitted by

GOKUL SRI RAM S - 921319104044

NARASH KUMAR A - 921319104131

EDWIN RAJAN G - 921319104041

BALA VIGNESH K G - 921319104029

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING



PSNA COLLEGE OF ENGINEERING AND TECHNOLOGY,
(An Autonomous Institution Affiliated to Anna University, Chennai)

DINDIGUL-624622

APRIL 2023

ADAPTIVE LEARNING METHOD FOR WEB RECOMMENDATION

A PROJECT REPORT

Submitted by

GOKUL SRI RAM S - 921319104044

NARASH KUMAR A - 921319104131

EDWIN RAJAN G - 921319104041

BALA VIGNESH K G - 921319104029

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING



PSNA COLLEGE OF ENGINEERING AND TECHNOLOGY,
(An Autonomous Institution Affiliated to Anna University, Chennai)

DINDIGUL-624622

APRIL 2023

PSNA COLLEGE OF ENGINEERING AND TECHNOLOGY,
(An Autonomous Institution Affiliated to Anna University, Chennai)
DINDIGUL-624622

BONAFIDE CERTIFICATE

Certified that this project report “**ADAPTIVE LEARNING METHOD FOR WEB RECOMMENDATION**” is the bonafide work of “**GOKUL SRI RAM S (921319104044), NARASH KUMAR A (921319104131), EDWIN RAJAN G (921319104041), BALA VIGNESH K G (921319104029)** who carried out the project work under my supervision.

SIGNATURE

Dr. D. SHANTHI M.E., Ph.D.,

HEAD OF THE DEPARTMENT

Department of CSE,
PSNA College of Engineering and
Technology,
Dindigul - 624622.

SIGNATURE

**Dr. J. BENADICT RAJA M.E.,
Ph.D.,**

SUPERVISOR

Associate Professor
Department of CSE,
PSNA College of Engineering and
Technology,
Dindigul - 624622.

Submitted for the project Viva- Voce Examination held on.....

INTERNAL EXAMINER

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

With warm hearts and immense pleasure, I thank the Almighty for his grace and blessings which drove me to the successful completion of the project. I would like to express my gratitude towards my parents for their kind cooperation and encouragement which help me in completion of this project.

I take this opportunity to express my sincere thanks to the respected Chairperson **Tmt. K. DHANALAKSHMI AMMAL**, who is the guiding light for all the activities in my college. I would like to express my deep gratitude to our Pro Chairman **Rtn. Thiru R.S.K RAGURAAM D.A.E, M.COM** for their continuous support towards the development of the students.

I would like to thank our Principal **Dr. D. VASUDEVAN M.E., Ph.D.**, for being a beacon in guiding every one of us and infusing us with the strength and enthusiasm to work over successfully.

I express my sincere thanks and heartfelt gratitude to **Dr. D. SHANTHI M.E., Ph.D.**, Professor and Head, Department of Computer Science and Engineering for her valuable suggestions and continuous encouragement in the completion of the project work.

This project would not have been possible without the motivation and guidance of our project coordinator **Dr. M. S. THANABAL M.E., Ph.D.**, Professor and project guide **Dr. J. BENADICT RAJA M.E., Ph.D.**, Associate Professor of the department of Computer Science and Engineering.

ABSTRACT

Web Usage Mining plays an important role in recommender systems and web personalization. In this work, It is proposed an effective recommender system based on ontology and Web Usage Mining. The first step of the approach is extracting features from web documents and constructing relevant concepts. Then build ontology for the web site use the concepts and significant terms extracted from documents. It is an integrated system to support learners' reviews. The review dashboard is used to recommend review contents that are adaptive to the individual learner's level of understanding and to present other information that is useful for review. The pages of the digital learning materials that are estimated to be insufficiently understood by each learner and the web pages related to those pages are recommended. As a method for estimating such pages, It is considered extracting the pages related to the questions that were answered incorrectly. According to the semantic similarity of web documents to cluster them into different semantic themes, the different themes imply different preferences. The proposed approach integrates semantic knowledge into Web Usage Mining and personalization processes.

TABLE OF CONTENTS

CHAPTER NO	TITLE	PAGE NO
	ABSTARCT	iv
	LIST OF FIGURES	viii
	LIST OF TABLES	ix
	LIST OF ABBREVATIONS	x
1.	INTRODUCTION	1
	1.1. PROJECT OVERVIEW	1
	1.1.1. Google	2
	1.1.2. Yahoo	3
	1.1.3. Duck Duck Go	4
	1.1.4. Bing	5
	1.1.5. Ask.com	6
	1.1.6. Google Patent	7
	1.2. OBJECTIVE	8
2.	LITERATURE SURVEY	9
	2.1. EXISISTING SYSTEM	15
	2.1.1. Disadvantages	16
	2.2. PROPOSED SYSTEM	17
	2.2.1. Advantages	17
3.	PROPOSED SYSTEM	18
	3.1. INTRODUCTION	18
	3.2. SYSTEM ARCHITECTURE	19

3.3. MODULE DSECRPTION	20
3.3.1. Creating Search History	20
3.3.2. Query Clustering	21
3.3.3. Query Reformulation	21
3.3.4. History Grouping	21
3.4. PSEUDO DOCUMENT	22
3.5. USER SEARCH GOALS	23
3.6. FRAME WORK OF OUR APPROACH	23
3.7. FEEDBACK SESSIONS	24
3.8. METHODOLOGY	24
3.8.1. Waterfall Approach	24
3.8.2. Analysis	25
3.8.3. Design	26
3.8.4. Coding	26
3.8.5. Testing	26
3.8.6. Integration	26
3.9. UML DIAGRAMS	26
3.10. CLASS DIAGRAM	28
3.11. DATA FLOW DIAGRAM	28
3.11.1. Level-0	29
3.11.2. Level-1	30
3.11.3. Level-2	30
3.12. USE CASE DIAGRAM	31
3.13. SEQUENCE DIAGRAM	32
3.14. COLLABORATION DIAGRAM	33
3.15. ACTIVITY DIAGRAM	34

4.	IMPLEMENTATION AND TESTING	36
	4.1. H/W SYSTEM CONFIGURATION	36
	4.2. S/W SYSTEMCONFIGURATION	36
	4.3. SOFTWARE ENVIRONMENT	36
	4.3.1. Java Programming Language	36
	4.3.2. JavaServer Pages	37
	4.3.3. Servlets	37
	4.3.4. MySQL	38
	4.3.5. Stemming Algorithm	38
	4.4. IMPLEMENTATION	39
	4.5. SYETM TESTING	43
	4.6. TYPES OF TESTING	44
	4.7. TEST CASES	46
	4.8. RESULTS AND DISCUSSION	46
	4.8.1. Result 1	46
	4.8.2. Result 2	47
5.	CONCLUSION	48

LIST OF FIGURES

FIGURE NO	FIGURE NAME	PAGE NO
3.1	System Architecture	19
3.2	Module description	20
3.3	Pseudo Documents	22
3.4	User Search Goals	23
3.5	Waterfall Model	25
3.6	Class Diagram	28
3.7	Level-0	29
3.8	Level-1	30
3.9	Level-2	31
3.10	Use Case Diagram	32
3.11	Sequence Diagram	33
3.12	Collaboration Diagram	34
3.13	Activity Diagram	35
4.1.	Registration Page	39
4.2.	Login Page	40
4.3.	Search Page	40
4.4.	Result Page	41
4.5.	Selected Link	41
4.6.	Clickthrough Submit	42
4.7.	Submitted Clickthrough	42
4.8.	Recommended Search Result	43
4.9.	Result – False	46
4.10.	Clickthrough Submit	47
4.11.	Result – True	47

LIST OF TABLES

TABLE NO	TABLE NAME	PAGE NO
4.1.	Test Cases	46

LIST OF ABBREVIATIONS

KT	Knowledge Tracing
KQN	Knowledge Query Network
RF	Random Forest
UML	Unified Modelling Language
DFD	Data Flow Diagram
JSP	JavaServer Pages
JSF	JavaServer Faces
SQL	Structured Query Language
WUM	Web Usage Mining
OOP	Object-oriented programming

CHAPTER 1

INTRODUCTION

1.1. PROJECT OVERVIEW

Web search engines are designed to fulfill user's information needs by receiving queries that represent their search intentions. However, queries can often be ambiguous and cover broad topics, leading to different users having diverse information needs even if they submit the same query. This can result in search engines providing irrelevant or inadequate results, frustrating users and leading to a poor user experience.

To address this issue, capturing user search goals can greatly enhance search engine relevance and user experience. User search goals refer to the different aspects of a query that user groups want to obtain. These goals can be considered as clusters of information needs for a query. By understanding and analysing these clusters, search engines can provide more relevant and targeted results that satisfy users' search intentions.

Analysing user search goals has numerous advantages for improving search engine performance and user experience. One such advantage is the ability to restructure search results based on user search goals. This involves grouping search results according to the search goals they fulfill, allowing users to easily find what they are looking for. This can enhance search engine relevance and efficiency, as well as making the search process more intuitive and user-friendly.

Another advantage of analysing user search goals is the ability to improve query recommendations. User search goals represented by some keywords can be utilized in query recommendation, suggesting more precise and relevant queries that better capture users' search intentions. This can save users time and effort, as well as improve the accuracy and relevance of search results.

Finally, analysing user search goals can also be useful in enhancing re-ranking of search results. The distributions of user search goals can be used to identify which search results are more relevant to which goals. By re-ranking the search results based on user search goals, the search engine can provide more accurate and targeted results that better satisfy users' search intentions.

Overall, analysing user search goals can greatly enhance search engine relevance and user experience. By understanding users' search intentions and clustering them into meaningful search goals, search engines can provide more targeted and accurate results, improving user satisfaction and engagement.

1.1.1. Google

Google Search Engine is a widely used web-based search engine that allows users to search for information on the internet. It was launched in 1997 by Google LLC, a subsidiary of Alphabet Inc. Google Search Engine uses sophisticated algorithms to crawl, index, and rank billions of web pages from across the internet, providing users with highly relevant search results based on their queries.

Google Search Engine features a simple and user-friendly interface where users can enter their search queries and receive a list of search results. The search results are ranked based on relevance, with the most relevant results appearing at the top of the list. Users can click on the search results to access the web pages containing the information they are looking for. In addition to basic web search functionality, Google Search Engine offers a wide range of features and services, such as image search, news search, video search, maps, and shopping search. It also provides specialized search results for specific industries, such as flights, hotels, and local businesses, as well as other tools like calculator, weather information, and language translations.

Google Search Engine is known for its advanced search capabilities, including the use of natural language processing, machine learning, and personalized search results based on user preferences and search history. It also offers search filters, such as time range, location, and file type, to help users refine their search queries and find more specific information.

Google Search Engine has continuously evolved and improved over the years, with regular updates to its search algorithms and user interface. It has become the dominant search engine globally, with a significant market share in many countries. As of September 2021, Google's global search market share was around 92%, according to StatCounter, making it the most widely used search engine worldwide.

1.1.2. Yahoo

Yahoo Search is a web search engine owned by Yahoo, a subsidiary of Verizon Communications. It was first launched in 1995 and has since become one of the most popular search engines on the web.

The Yahoo search engine uses a proprietary algorithm to crawl and index web pages. This algorithm determines which pages are most relevant to a user's search query, and presents them in the search results. Yahoo also uses a variety of other factors to determine page ranking, including page quality, keyword density, and inbound links.

The Yahoo search engine uses a proprietary algorithm to crawl and index web pages. This algorithm determines which pages are most relevant to a user's search query, and presents them in the search results. Yahoo also uses a variety of other factors to determine page ranking, including page quality, keyword density, and inbound links.

Yahoo has also integrated social media features into its search engine, allowing users to view tweets, photos, and other content from social media sites directly in the search results. This feature is called Yahoo Social Bar.

Overall, Yahoo Search has been a popular and reliable search engine for many years, offering a range of features and services to help users find the information they need on the web.

1.1.3. Duck Duck Go

DuckDuckGo is a privacy-focused web search engine that allows users to search for information on the internet while prioritizing user privacy and data protection. It was launched in 2008 by DuckDuckGo Inc., a company headquartered in Pennsylvania, United States. DuckDuckGo differentiates itself from other search engines by not tracking user search queries or personal information, and by delivering search results without personalized ads.

One of the main features of DuckDuckGo is its emphasis on user privacy. DuckDuckGo does not store any search history or track user activity, ensuring that users' search queries and personal information remain private. DuckDuckGo also uses encryption (HTTPS) for all search queries and results, providing an added layer of security.

DuckDuckGo's search results are generated from multiple sources, including its own web crawler as well as results from other search engines, such as Bing and Yahoo. DuckDuckGo's search results are presented in a clean and minimalist interface, with a focus on providing relevant and unbiased search results without personalized ads or tracking.

DuckDuckGo also offers additional features, such as "Instant Answers," which provides quick answers to common questions directly in the search results page, and "Bangs," which allow users to search specific websites or perform actions directly from the DuckDuckGo search bar using shortcuts.

DuckDuckGo has gained popularity among users who prioritize privacy and data protection in their online activities. It has also been endorsed by privacy advocates and organizations for its commitment to user privacy. However, as of September 2021, DuckDuckGo's global search market share remains relatively small compared to larger search engines like Google, with around 2% of the global search market share, according to StatCounter.

1.1.4. Bing

Bing is a web search engine developed and operated by Microsoft. It was launched in 2009 as a replacement for Microsoft's earlier search engine, Live Search. Bing aims to provide users with a visually appealing and intuitive search experience, with a focus on delivering relevant and comprehensive search results.

Bing uses a combination of web crawling, indexing, and ranking algorithms to generate search results. Bing's web crawler, known as Bing Bot, continuously crawls the internet to discover and index web pages. The indexed web pages are then ranked by Bing's algorithms based on a variety of factors, including relevance, authority, and user engagement signals, to determine the most relevant search results for a given query. One of Bing's notable features is its visually rich search results page, which includes features such as a daily changing background image, video previews, and interactive features like image and video carousels. Bing also offers a "Snapshot" feature that provides additional information related to a search query, such as related people, news articles, and events, directly in the search results page.

Bing also offers specialized search results for specific topics, such as images, videos, news, maps, and shopping. Bing's image search, for example, allows users to search for images with specific sizes, colours, and layouts, and offers features such as image search trends and image insights to help users find relevant images.

Bing also powers the search results for other Microsoft products, such as Microsoft Edge browser, Cortana virtual assistant, and Microsoft Office suite. Bing also has partnerships with other websites and search engines to provide search results for their platforms.

Bing has gained a significant share of the search market, particularly in some regions, although it still lags behind Google in terms of global search market share. Bing continues to evolve and innovate, with ongoing updates to its search algorithms, features, and user interface to provide users with an enhanced search experience.

1.1.5. Ask.com

Ask.com is a search engine that is known for its question-and-answer format. It provides answers to users' queries based on its extensive database of question-and-answer pairs. Users can submit questions in plain English or keyword-based queries, and Ask.com will attempt to provide relevant answers.

Ask.com also offers a variety of other features, including image search, news search, and video search. Users can also customize their search settings, such as language preferences and search filters. One notable feature of Ask.com is its "Smart Answers" feature, which provides direct answers to certain types of queries, such as weather forecasts, currency conversions, and time zone information, without having to click on search results.

Ask.com also provides a search toolbar that can be installed on web browsers for easy access to its search functionality. The search engine has a user-friendly interface with a focus on simplicity and ease of use.

However, it's worth noting that Ask.com has a smaller market share compared to some of the more popular search engines like Google and Bing, and its search results may not be as comprehensive or accurate in certain cases. Nevertheless, Ask.com can be a useful option for users who prefer a question-and-answer format and are looking for quick answers to specific queries.

1.1.6. Google Patent

Google Patent Search is a specialized search engine provided by Google that focuses on searching and retrieving patents from around the world. It allows users to search for patents based on keywords, patent numbers, inventor names, assignee names, and other criteria related to patents.

Google Patent Search provides access to a vast database of patents from various patent offices, including the United States Patent and Trademark Office (USPTO), the European Patent Office (EPO), the World Intellectual Property Organization (WIPO), and many others. Users can search for patents by country, patent type, filing date, and other parameters to narrow down their search results.

The search results in Google Patent Search typically include patent titles, abstracts, drawings, and other relevant information about the patents. Users can view the full text of the patents, including the detailed description and claims, in PDF format. One notable feature of Google Patent Search is its ability to highlight relevant terms in the patent documents, making it easier for users to quickly identify the relevant information they are looking for.

1.2. OBJECTIVE

The proposed system aims to support learners in their review of digital learning materials by providing personalized recommendations of review contents. The system uses a review dashboard that recommends adaptive review contents based on each individual learner's level of understanding. The review dashboard also presents other information that is useful for review.

To generate personalized recommendations, the system identifies the pages of digital learning materials that each learner is estimated to have insufficient understanding of, as well as the web pages related to those pages. One method used to estimate such pages is by extracting the pages related to questions that were answered incorrectly by the learner.

To support the recommendation process, the system utilizes Web Usage Mining and ontology-based techniques. The first step in the approach is to extract features from web documents and construct relevant concepts. These concepts are then used to build an ontology for the web site. The ontology is based on the significant terms and concepts extracted from the web documents.

The system also clusters web documents into different semantic themes based on their semantic similarity. Different themes imply different preferences, which are used to personalize the recommendations further. The proposed approach integrates semantic knowledge into the Web Usage Mining and personalization processes.

CHAPTER 2

LITERATURE SURVEY

A. Title: Developing an early-warning system for spotting at-risk students by using e-book interaction logs.

Author: G. Akc,apınar, M. N. Hasnine, R. Majumdar, B. Flanagan, and H. Ogata

Technology Used: J48 algorithm

Description:

This work describes the development of an early warning system for identifying students who are at risk of academic failure in the context of eBook-based teaching and learning. The system is designed to analyse students' eBook reading data to predict those who may be struggling with the course material and in danger of falling behind. To create the system, the authors used 13 prediction algorithms, including the J48 algorithm, to create prediction models using data from different weeks of the course.

The goal of using multiple prediction algorithms was to determine which algorithm provided the most accurate predictions for identifying at-risk students. By testing each algorithm with the same data set, the authors were able to compare the performance of each model and select the best performing algorithm. Once the best algorithm was identified, the authors were able to use it to determine the optimum time for possible interventions to support at-risk students.

By analysing eBook interaction logs, the early warning system can provide educators with valuable insights into how students are engaging with the course material. This information can be used to identify patterns of behaviour that may indicate a student is struggling and may benefit from additional support. By intervening early, educators can provide targeted interventions to help these students succeed in their academic pursuits.

Overall, the development of this early warning system demonstrates the potential of using machine learning algorithms and eBook interaction logs to improve educational outcomes. By leveraging these technologies, educators can gain valuable insights into how students are engaging with the course material and provide them with timely support to help them succeed.

B. Title: An e-learning recommendation approach based on the self-organization of learning resource

Author: S. Wan and Z. Niu

Technology Used: Combining learner influence model (LIM)

Description:

Recommender systems are commonly used in e-learning platforms to provide personalized recommendations to learners, such as recommending relevant learning resources, courses, or activities. However, in some cases, interpersonal information between learners is scarce, which makes it difficult to apply collaborative filtering techniques to achieve recommendations. To address this issue, a proposed approach is to simulate learning objects (LOs) as intelligent entities using the self-organization theory.

The self-organization theory is a concept in artificial intelligence that enables entities to self-organize, adapt, and learn based on their interactions with their environment. In this approach, LOs are simulated as intelligent entities that can receive, transmit, and move based on their interactions with the environment and other LOs. LOs can also store and process information, such as learner preferences, which can be used to generate recommendations.

The first step in this approach is to simulate LOs as intelligent entities. This involves defining the behaviour and interactions of LOs with each other and with the environment. For example, LOs may move towards other LOs with similar content, or may transmit information about learner preferences to other LOs. This

behaviour is based on the self-organization theory, which allows the LOs to adapt and learn based on their interactions with the environment and other LOs.

The second step is to design an environment perception module. This module can capture and perceive learner's preference drifts by analysing the self-organization behaviours of LOs. The environment perception module can use machine learning algorithms to analyse the interactions between LOs and the environment, and identify patterns in learner preferences.

The final step is to generate recommendations based on the learners' explicit requirements and implicit preference drifts. Recommendations are generated through LOs' self-organization behaviours, which are based on the preferences and interactions of other learners. LOs can use this information to generate recommendations that are personalized and relevant to each learner.

This approach has been applied to real-life learning processes, and ample experimental results demonstrate the high adaptability, diversity, and personalization of the recommendations. The approach is particularly useful in cases where interpersonal information is scarce, such as in online learning environments with a large number of learners. By simulating LOs as intelligent entities, this approach enables personalized and relevant recommendations to be generated even when there is limited interpersonal information available.

C. Title: Knowledge query network for knowledge tracing: How knowledge interacts with skills

Author: J. Lee and D.-Y. Yeung

Technology Used: Knowledge Tracing (KT)

Description:

This work discusses the challenges of modelling students' knowledge states and skills in a way that accurately captures the interactions between these two factors. To address this challenge, the authors propose a new model called the Knowledge Query Network (KQN).

KQN uses neural networks to encode student learning activities into knowledge state and skill vectors. These vectors capture the student's level of knowledge and skills in various areas, and are updated as the student engages in further learning activities.

One of the key features of KQN is its ability to model the interactions between knowledge state and skill vectors using the dot product. This allows the model to capture the complex relationships between these two factors and generate more accurate predictions of students' learning outcomes.

Another innovative aspect of KQN is its use of probabilistic skill similarity, which relates the pairwise cosine and Euclidean distances between skill vectors to the odds ratios of the corresponding skills. This makes KQN more interpretable and intuitive, as it provides a way to understand how different skills are related to each other and how they contribute to students' overall knowledge state.

Overall, KQN represents a significant advance in the field of educational modelling, as it provides a more accurate and nuanced way to capture the complex interactions between students' knowledge states and skills. By using neural networks and probabilistic skill similarity, KQN can generate more personalized and effective recommendations for students, and help educators better understand and support students' learning processes.

D. Title: Recommending remedial readings using student knowledge state.

Author: K. Thaker, L. Zhang, D. He, and P. Brusilovsky

Technology Used: Random Forest (RF)

Description:

The work aims to automatically identify textbook materials that are most relevant and suitable to students, with a focus on incorporating their current knowledge state on domain concepts associated with the activity to recommend personalized remedial sections. The following is a detailed explanation of the methods and experiments conducted in this work.

The first step is to identify the domain concepts associated with the activity, which are the key topics covered in the textbooks. This is done by analysing the content of the textbooks and identifying the main topics covered. These domain concepts are used to represent the knowledge space of the domain.

The next step is to assess the student's knowledge state on each of these domain concepts. This is done by administering a pre-test to assess the student's baseline knowledge and then tracking their progress over time using various assessment tools, such as quizzes and exams. This information is used to generate a knowledge profile for each student, which represents their current knowledge state on each domain concept.

Using the knowledge profiles, the system can then recommend remedial sections that are personalized to each student. This is done by first identifying the relevant textbook sections that cover the domain concepts in which the student is struggling. The system then selects the remedial sections that are most relevant and suitable to the student based on their knowledge profile.

To evaluate the effectiveness of this approach, experiments were conducted on a sample of students. The system was able to accurately identify the remedial sections that were most relevant and suitable to each student based on their knowledge profile. The students who received the personalized remedial sections showed significant improvement in their knowledge state on the relevant domain concepts compared to those who did not receive personalized recommendations.

In conclusion, this work provides a method for automatically identifying the textbook materials that are most relevant and suitable to students based on their knowledge state on domain concepts. By incorporating the student's current knowledge state, the system can recommend personalized remedial sections to improve their understanding of the key topics covered in the textbooks. The experiments conducted demonstrate the effectiveness of this approach in improving student learning outcomes.

E. Title: Utilizing learners' negative ratings in semantic content-based recommender system for elearning forum

Author: N. A. Albatayneh, K. I. Ghauth, and F.-F. Chua

Technology Used: Semantic content-based filtering technique (SCB).

Description:

This work addresses the challenge of information overload in e-learning online discussion forums, where the amount of shared information can make it difficult for learners to identify interesting or relevant messages. To address this issue, the authors propose a recommendation architecture that utilizes a semantic content-based filtering approach and learners' negative ratings.

The architecture works as follows: first, the system extracts the semantic content of each post message using natural language processing techniques. Then, the system applies a content-based filtering approach to match the semantic content of each post message with learners' interests and preferences.

This filtering approach assumes that learners are interested in messages that are like the ones they have interacted with positively in the past.

In addition, the system also considers learners' negative ratings of post messages. This helps the system avoid recommending messages that learners have already found uninteresting or irrelevant in the past. The system uses collaborative filtering techniques to identify similar learners and takes their negative ratings into account when generating recommendations.

To evaluate the proposed recommender system, the authors compared its recommendation accuracy and its impact on learners' performance with existing e-learning recommender systems that use similar filtering techniques. The results showed that the proposed system outperformed the existing systems in terms of recommendation accuracy and had a positive impact on learners' performance. The authors suggest that their approach can be applied to other e-learning platforms to help learners discover relevant and interesting information more effectively.

2.1. EXISTING SYSTEM

Useful knowledge discovery from Web usage data and satisfactory knowledge representation for effective Web-page recommendations are crucial and challenging.

Existing system provide method to efficiently provide better Web-page recommendation through semantic enhancement by integrating the domain and Web usage knowledge of a website. Two new models are proposed to represent the domain knowledge.

The first model uses ontology to represent the domain knowledge. The second model uses one automatically generated semantic network to represent domain terms, Web-pages and the relations between them. Another new model, the conceptual prediction model, is proposed to automatically generate a semantic network of the semantic Web usage knowledge, which is the integration of domain knowledge and Web usage knowledge.

A number of queries have been developed to query about these knowledge bases. Based on these queries, a set of recommendation strategies have been proposed to generate Web-page candidates. The recommendation results have been compared with the results obtained from an advanced existing Web Usage Mining (WUM) method.

2.1.1. Disadvantages

Existing recommendation systems are: cold-start, sparsely, overspecialization and domain-dependency.

The performance of existing system depends on the sizes of training datasets. The bigger the training dataset size is, predicted pages are limited within the discovered Web access sequences.

The domain ontology can be constructed manually by experts, or by automatically learning models is need to design and implement the learning models which can only be done by professionals at the beginning.

2.2. PROPOSED SYSTEM

In proposed system present a personalized-recommendation system, a system that makes use of representations of items and user-profiles based on ontologies in order to provide semantic applications with personalized services.

The semantics method achieved by using two different methods. A domain-based method makes inferences about user's interests and a taxonomy-based similarity method is used to refine the item-user matching algorithm, improving overall results. The recommender proposed is domain-independent, is implemented as a Web service, and uses both explicit and implicit feedback-collection methods to obtain information on user's interests.

Proposed recommender system based on ontology and Web Usage Mining. The first step of the approach is extracting features from web documents and constructing relevant concepts. Then build ontology for the web site use the concepts and significant terms extracted from documents. According to the semantic similarity of web documents to cluster them into different semantic themes, the different themes imply different preferences.

2.2.1. Advantages

Integrating domain knowledge with Web usage knowledge enhances the performance of recommender systems using ontology-based Web mining techniques. The construction of this model is semi-automated so that the development efforts from developers can be reduced.

The user-profile learning algorithm, responsible for expanding and maintaining up-to-date the long-term user's interests, employs a domain-based inference method in combination with other relevance feedback methods to populate more quickly the user profile and therefore reduce the typical cold-start problem.

CHAPTER 3

PROPOSED SYSTEM

3.1. INTRODUCTION

This work proposes an effective recommender system that utilizes Web Usage Mining and ontology for web personalization. Web Usage Mining is a critical component of recommender systems, and this approach aims to improve learners' reviews. The first step is to extract relevant features from web documents and construct corresponding concepts. An ontology is then built for the website using the extracted concepts and significant terms from the documents. This system is an integrated one that supports learners' reviews by recommending review content adaptive to the individual learner's level of understanding and presenting other useful information.

The system recommends pages of digital learning materials estimated to be insufficiently understood by each learner and web pages related to those pages. The method for estimating such pages involves extracting pages related to the questions that were answered incorrectly. By clustering web documents into different semantic themes based on their semantic similarity, the system can provide personalized recommendations to users with different preferences. The proposed approach integrates semantic knowledge into Web Usage Mining and personalization processes, making it a powerful tool for enhancing web personalization and improving user experience.

The system architecture of the proposed system architecture is explained below in Figure 3.1

3.2. SYSTEM ARCHITECTURE

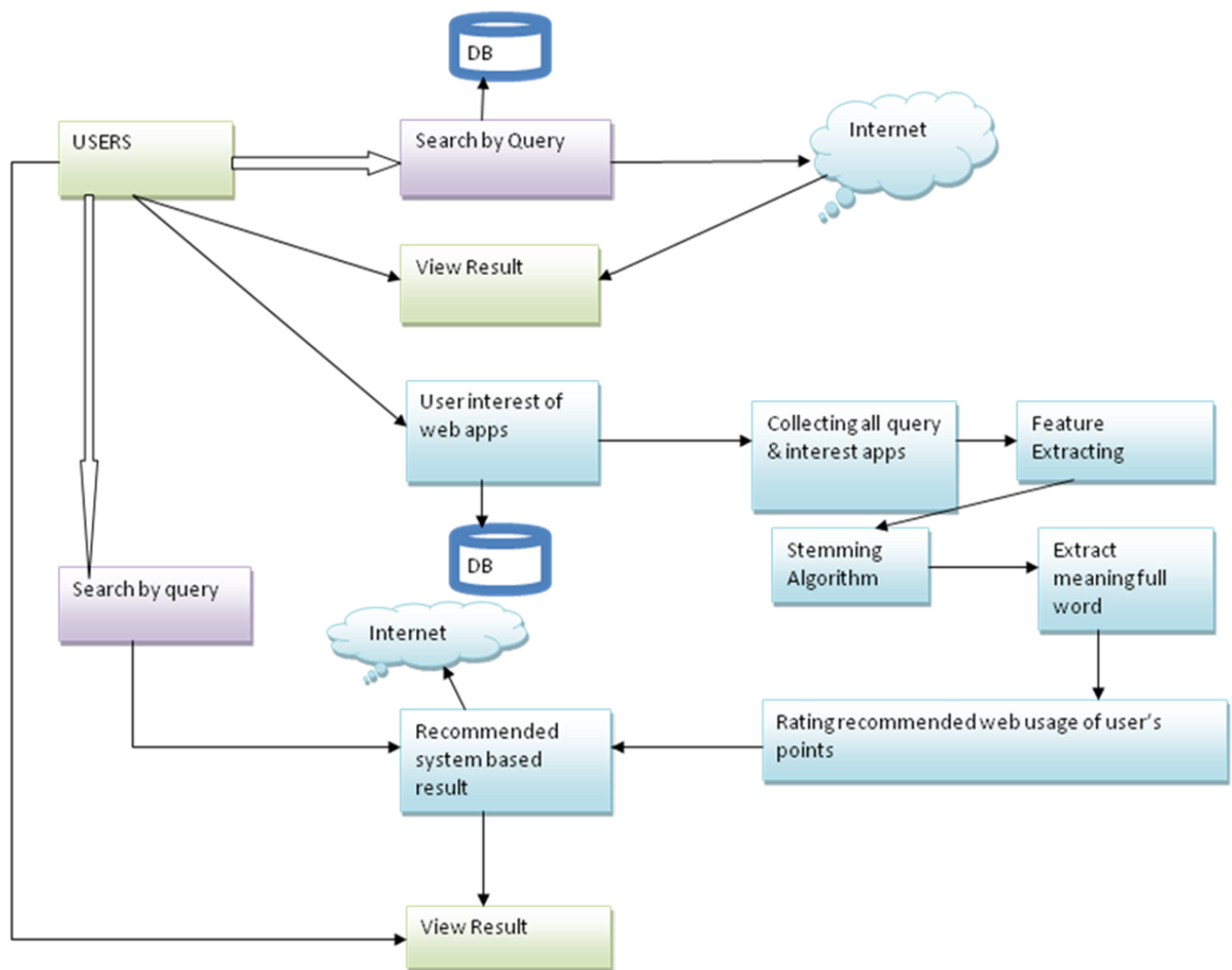


Figure 3.1. System Architecture

The user needs to log in with their login credentials. After logging in, they can enter a query and the system will display the corresponding search results page. From there, the user can click on a link to access more information and choose the "click through submit" option. Once submitted, the user can repeat the same search query to retrieve additional related queries and interesting applications. The results of this subsequent search will then be presented to the user.

3.3. MODULE DESCRIPTION

- Creating Search history
- Query clustering
- Query reformulation
- History grouping

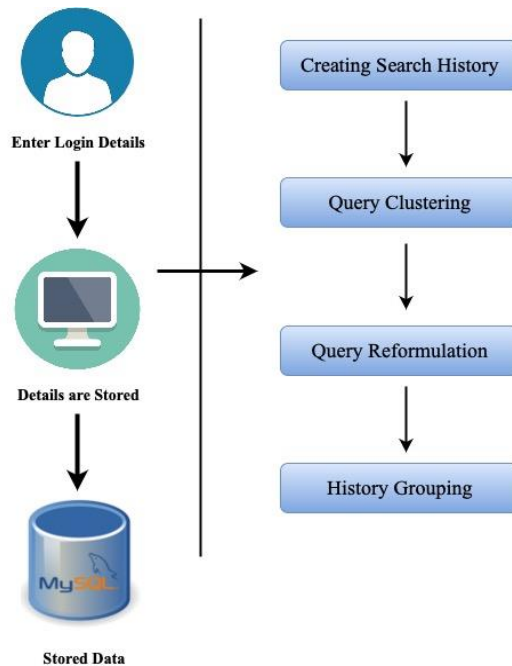


Figure 3.2. Module Description

3.3.1. Creating Search History

The data source for creating user profiles can include personal documents such as browsing history and emails stored on a user's computer. By focusing on frequently occurring terms within these documents (as shown in Figure 3.2.), the dimensionality of the document set is reduced, enabling a clearer understanding of the user's interests. This module enables search engines to gain a deeper understanding of a user's session and potentially customize their search experience based on their specific needs. After identifying query groups, search engines can accurately interpret the search context behind the current query by analysing queries and clicks within the corresponding group.

3.3.2. Query Clustering

To achieve personalization, users' queries are grouped into various query clusters using concept-based user profiles in the clustering process (as shown in Figure 3.2.). The clustering process involves identifying the most similar pair of concept nodes, merging the most similar pair of query nodes, and continuing the process iteratively. Each user's individual query is considered a separate node with a unique user identifier. The grouping process is carried out dynamically by initially assigning the current query and clicks to a query group.

3.3.3. Query Reformulation

For effective grouping of queries and clicks, it is crucial to establish appropriate relevance between query groups. The assumption is that users tend to make similar queries and clicks within a brief timeframe. Analysing the search history of numerous users provides signals about query relevance, such as which queries are frequently issued in conjunction with each other. This highlights the connection between queries that typically result in clicks on similar URLs. The query reformulation graph and query click graph derived from search logs are instrumental in determining the relevance between queries or query groups in a user's search history (as shown in Figure 3.2.).

3.3.4. History Grouping

One approach to grouping queries is to initially consider every query in a user's history as a query group, and then merge these groups iteratively using a k-means algorithm (as shown in Figure 3.2.). However, this approach is not practical in our scenario for two reasons. Firstly, it could potentially disrupt a user's existing query groups, thus undoing their manual efforts to organize their search history. Secondly, it involves a high computational cost as a large number of query group similarity computations would need to be repeated for every new query.

3.4. PSEUDO DOCUMENT

The aim of this work is to connect feedback sessions with pseudo documents that signify the user's search objectives, as demonstrated in Figure 3.3. This process consists of two primary steps. Firstly, a brief text paragraph that includes the webpage's title and snippet is generated to represent the URLs in the feedback session. Various textual processing methods, such as lowercase conversion, stemming, and stop-word elimination, are employed on these text paragraphs.

The second step involves forming a pseudo-document based on the URL representations obtained in the first step. An optimization approach is suggested to create a feature representation of a feedback session that combines both clicked and unclicked URLs in the feedback session. This approach allows considering both the pages that the user clicked on and the pages that they did not choose, but which may still provide valuable information about their search objectives.

In general, this method helps to create more precise and representative pseudo-documents that reflect the user's search intentions and enhance the search engine's performance in providing relevant search results.

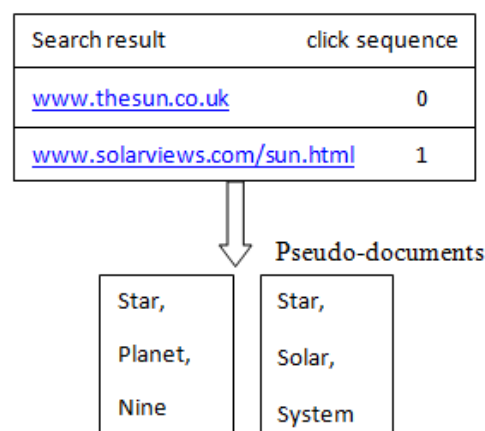


Figure 3.3. Pseudo Document

3.5. USER SEARCH GOALS

In this work pseudo-documents is clustered by K-means clustering which is simple and effective (as shown in Figure 3.4.). Since the exact number of user search goals for each query is unknown, K is set to five different values, and clustering is performed based on each value. After clustering all the pseudo-documents, each cluster is treated as a user search goal. The centre point of a cluster is computed as the average of the vectors of all the pseudo-documents in the cluster.

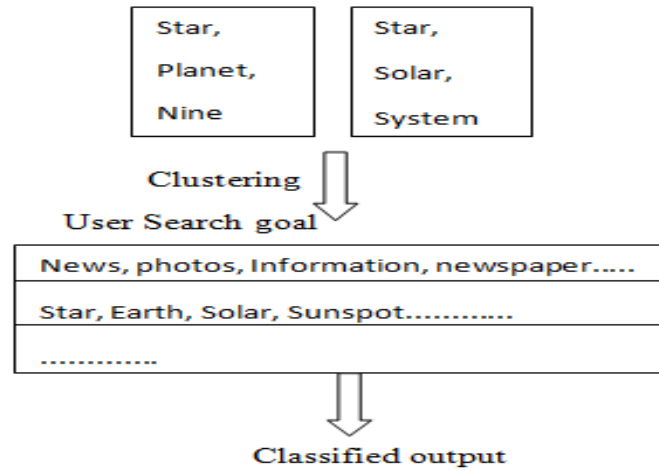


Figure 3.4. User Search Goals

3.6. FRAMEWORK OF OUR APPROACH

This proposed framework is divided into two parts by a dashed line. In the upper part, feedback sessions for a query are extracted from user click-through logs and mapped to pseudo-documents. User search goals are inferred by clustering these pseudo-documents and represented by key words. Since the exact number of user search goals is unknown, several different values are tried, and the optimal value is determined based on feedback from the bottom part. In the bottom part, the original search results are restructured based on the user search goals inferred from the upper part. The performance of restructuring the search results is evaluated using the proposed evaluation criterion CAP. The evaluation result is used as feedback to select the optimal number of user search goals in the upper part.

3.7. FEEDBACK SESSIONS

In this work, the focus is on inferring user search goals for a particular query. Therefore, a single session containing only one query is introduced, which distinguishes it from conventional sessions. Meanwhile, the feedback session in this work is based on a single session, although it can be extended to the whole session.

The proposed feedback session consists of both clicked and unclicked URLs and ends with the last URL that was clicked in a single session. The motivation is that before the last click, all the URLs have been scanned and evaluated by users. Therefore, besides the clicked URLs, the unclick done before the last click should be a part of the user feedbacks. From FIGURE 3.3. search results of the query “the sun” and the right part is a user’s click sequence where “0” means “unclicked.” Inside the feedback session, the clicked URLs tell what users require and the unclicked URLs reflect what users do not care about. It should be noted that the unclicked URLs after the last clicked URL should not be included in the feedback sessions since it is not certain whether they were scanned or not. Each feedback session can tell what a user requires and what they do not care about. Moreover, there are plenty of diverse feedback sessions in user click-through logs. Therefore, for inferring user search goals, it is more efficient to analyse the feedback sessions than to analyse the search results or clicked URLs directly.

3.8. METHODOLOGY

3.8.1. Waterfall Approach

The waterfall model is a software development methodology that follows a linear and sequential approach to project management. It involves a series of distinct stages, including requirements gathering and analysis, design, coding, testing, and integration (as shown in Figure 3.5.). Each stage must be completed

before the next one can begin, and changes to the product are difficult and costly to make once a stage is complete. While the model is useful for projects with well-defined requirements, it may not be suitable for projects that require flexibility or involve significant uncertainty.

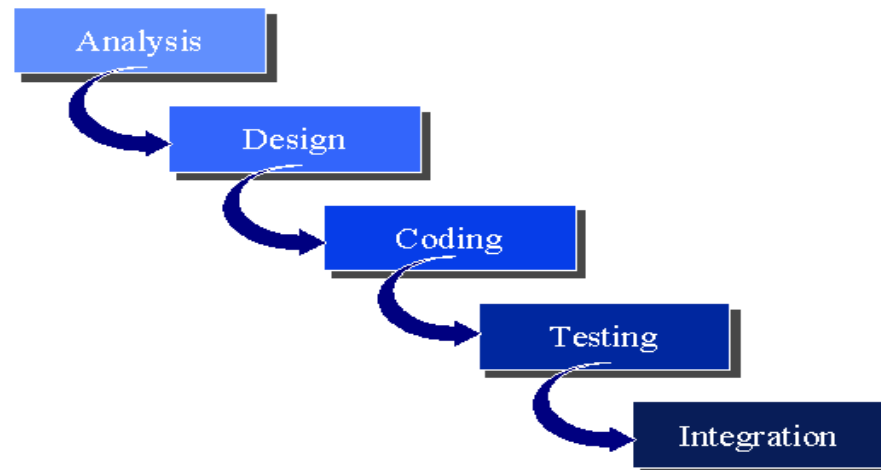


Figure 3.5. Waterfall Model

3.8.2. Analysis

Current recommendation systems face challenges such as cold-start, sparsity, overspecialization, and domain-dependency, with their performance relying on the size of the training datasets. Ontology construction can be done manually or with automated models but requires professional expertise during initial design and implementation.

The proposed personalized-recommendation system utilizes ontologies and Web Usage Mining to extract features from web documents, construct relevant concepts, and cluster them into different semantic themes to provide more accurate and personalized recommendations for semantic applications.

3.8.3. Design

To start, the user must log in using their credentials. Once logged in, they can search for a query, and the search results page will be shown. From there, the user can click on a link to view more details and select the "click through submit" option. After that, the user can perform the same search query to collect all related queries and interesting applications by Stemming Algorithm. The results of this subsequent search will then be displayed as shown in Figure 3.1.

3.8.4. Coding

The personalized recommender system employs Java for its backend, JSP for its front-end, and utilizes a MySQL database to store URL and profile information.

3.8.5. Testing

The personalized recommender system undergoes evaluation through both functional testing, which verifies that each function operates as intended, and performance testing, which measures the system's response time, stability, and scalability under various load conditions.

3.8.6. Integration

The individual software modules are combined and tested as a group to verify that they function correctly together.

3.9. UML DIAGRAMS

UML stands for Unified Modelling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object-oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modelling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modelling and other non-software systems. The UML represents a collection of best engineering practices that have proven successful in the modelling of large and complex systems.

The UML is a very important part of developing objects-oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

Goals

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modelling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modelling language.
5. Encourage the growth of OO tools market.

3.10. CLASS DIAGRAM

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information (as shown in Figure 3.6.).

The Proposed system is implemented with 4 different classes (UserClass, Search, WebPage Content, Onclick), the figure 3.6 shows the methods and interlink of various classes.

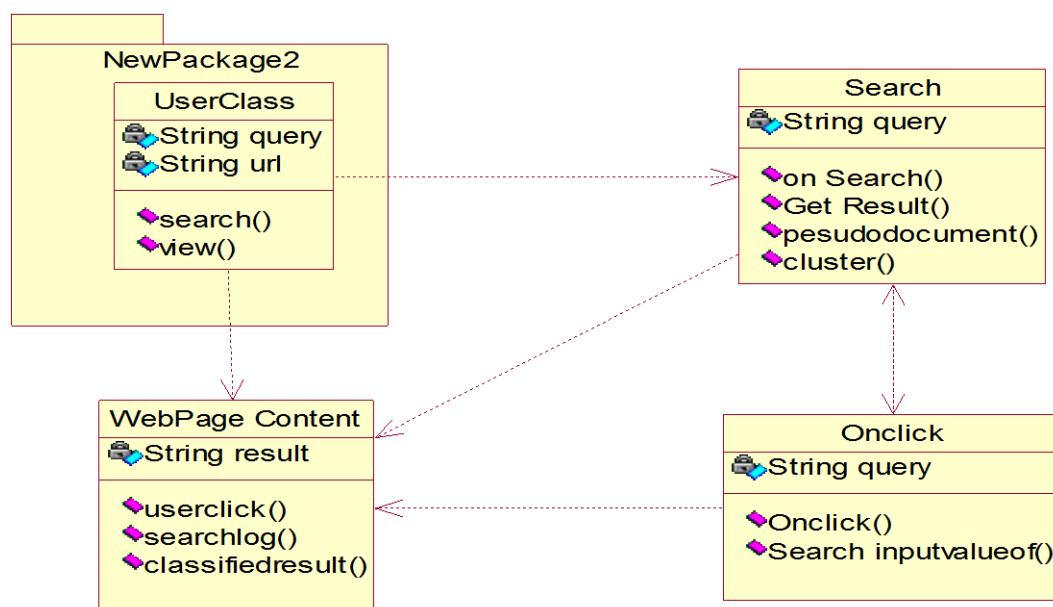


Figure 3.6. Class Diagram

3.11. DATA FLOW DIAGRAM

A Data Flow Diagram (DFD) is a visual representation of a system that shows how data moves between different components of the system. The purpose of a DFD is to model the flow of data through a system and to help understand, analyse, and communicate how the system works.

DFDs use a set of symbols and connectors to represent the different components of the system, including processes, data stores, data flows, and external entities. They can be developed at different levels of detail, ranging from high-level context diagrams to detailed diagrams showing individual processes and data flows.

3.11.1. Level-0

A level 0 Data Flow Diagram (DFD) is a high-level diagram that provides a broad overview of the system and its interaction with external entities (as shown in Figure 3.7.). It shows the main processes that make up the system and the data flows between them.

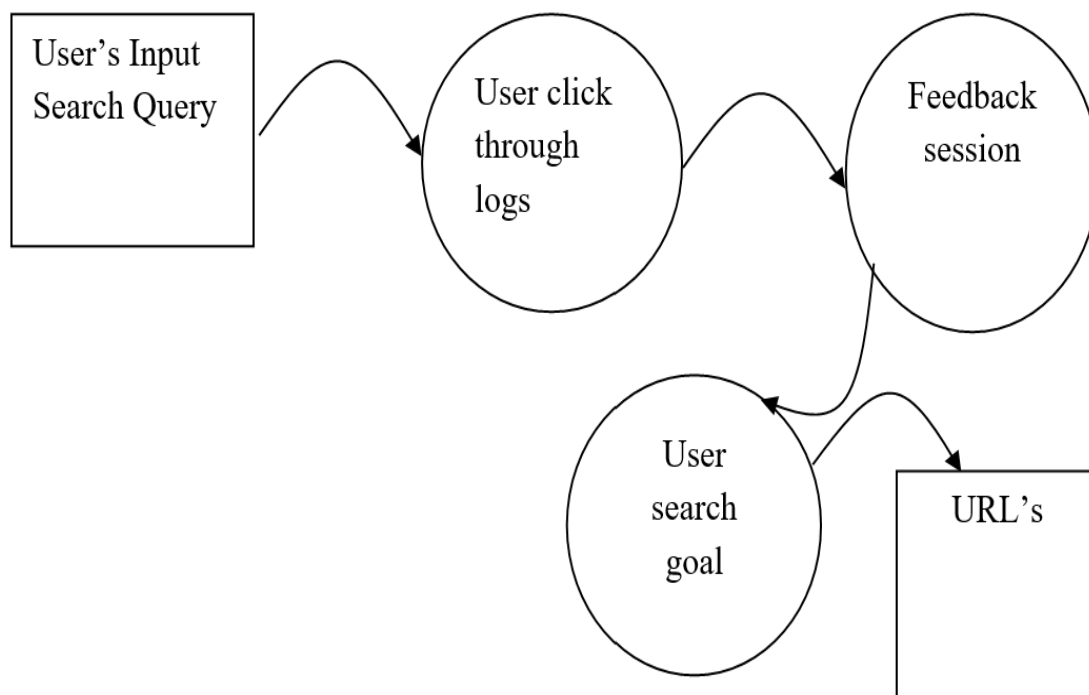


Figure 3.7. Level-0

3.11.2. Level-1

A level 1 Data Flow Diagram (DFD) provides a more detailed view of the system than the level 0 DFD (as shown in Figure 3.8.). It shows the main processes and sub-processes of the system, as well as the data stores that they interact with.

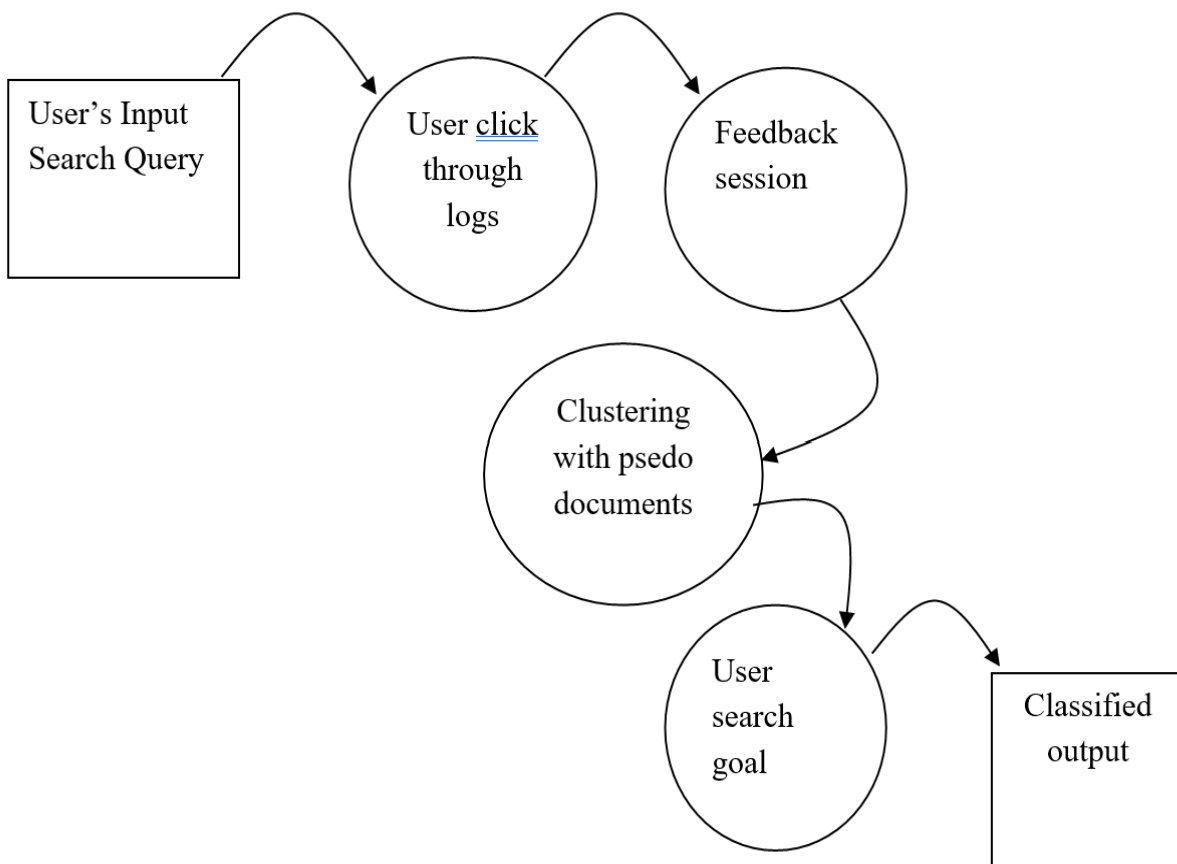


Figure 3.8. Level-1

3.11.3. Level-2

A level 2 Data Flow Diagram (DFD) provides even more detail than the level 1 DFD (as shown in Figure 3.9.). It breaks down the sub-processes identified in the level 1 DFD into further detail and shows the flow of data within each process

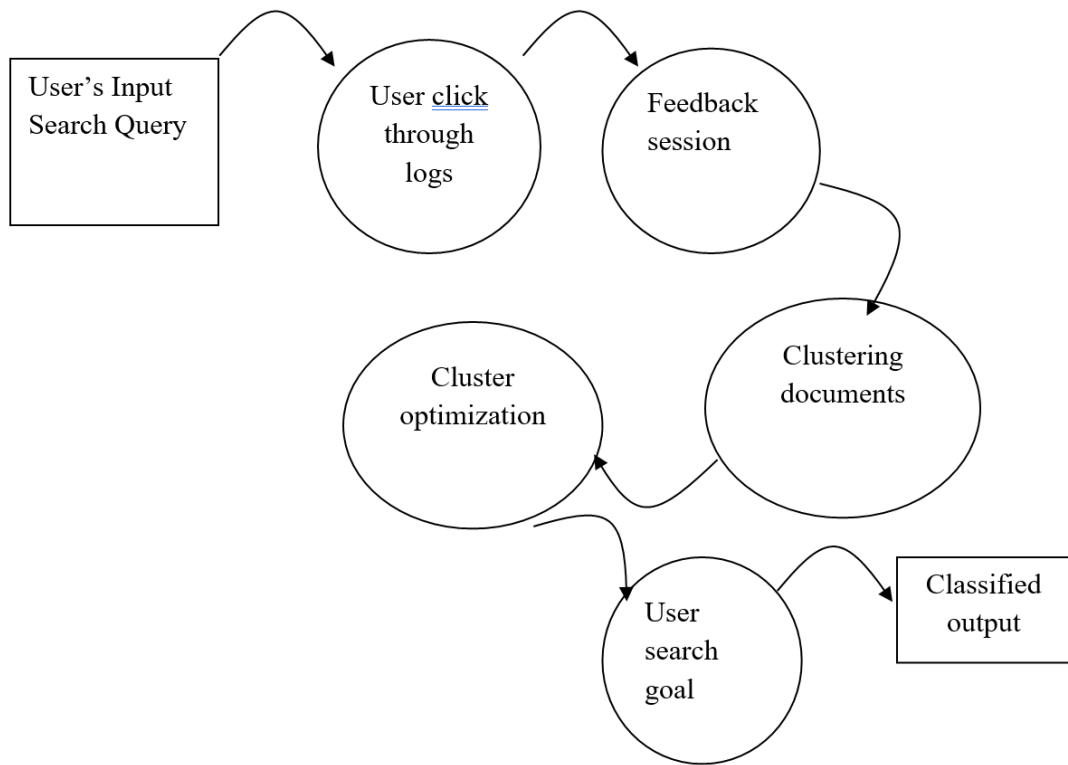


Figure 3.9. Level-2

3.12. USE CASE DIAGRAM

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis (as shown in Figure 3.10.). Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases.

The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

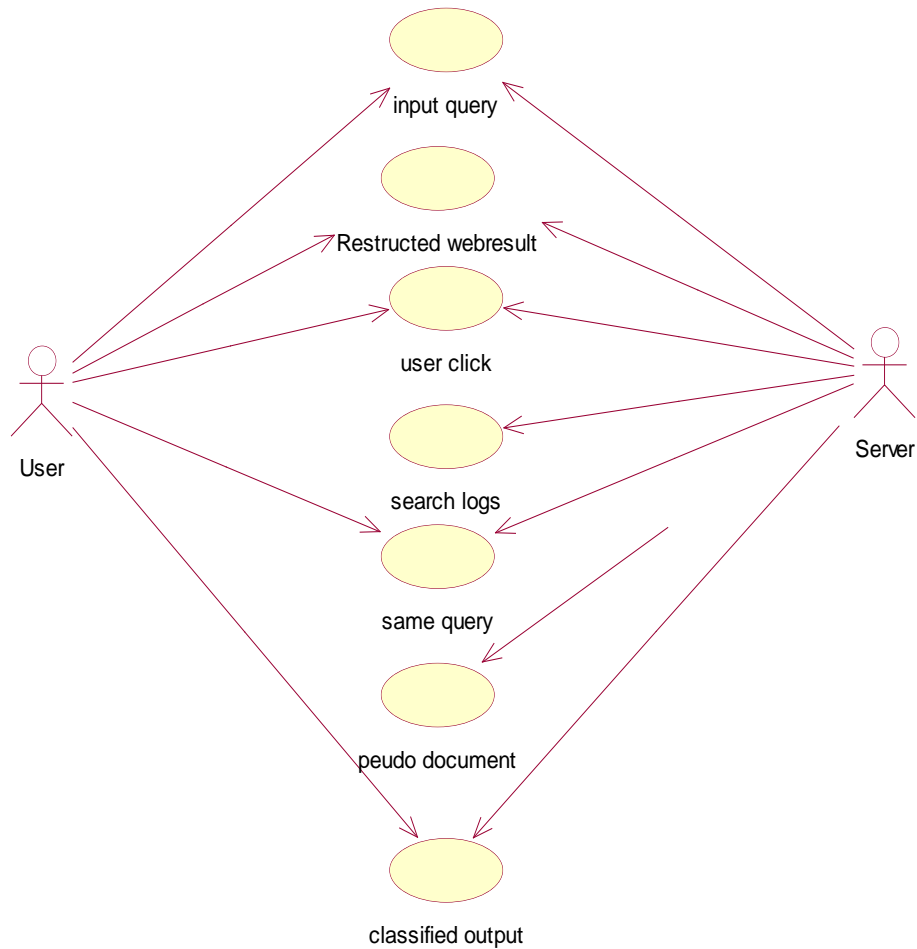


Figure 3.10. Use Case Diagram

3.13. SEQUENCE DIAGRAM

A sequence diagram is a type of UML (Unified Modelling Language) diagram that shows the interactions and messages exchanged between objects or components in a system over time. It provides a dynamic view of how objects collaborate to fulfill a specific behaviour or achieve a certain task (as shown in Figure 3.11.). Sequence diagrams are commonly used in software development to model the flow of interactions between different objects or components during the execution of a use case or a scenario.

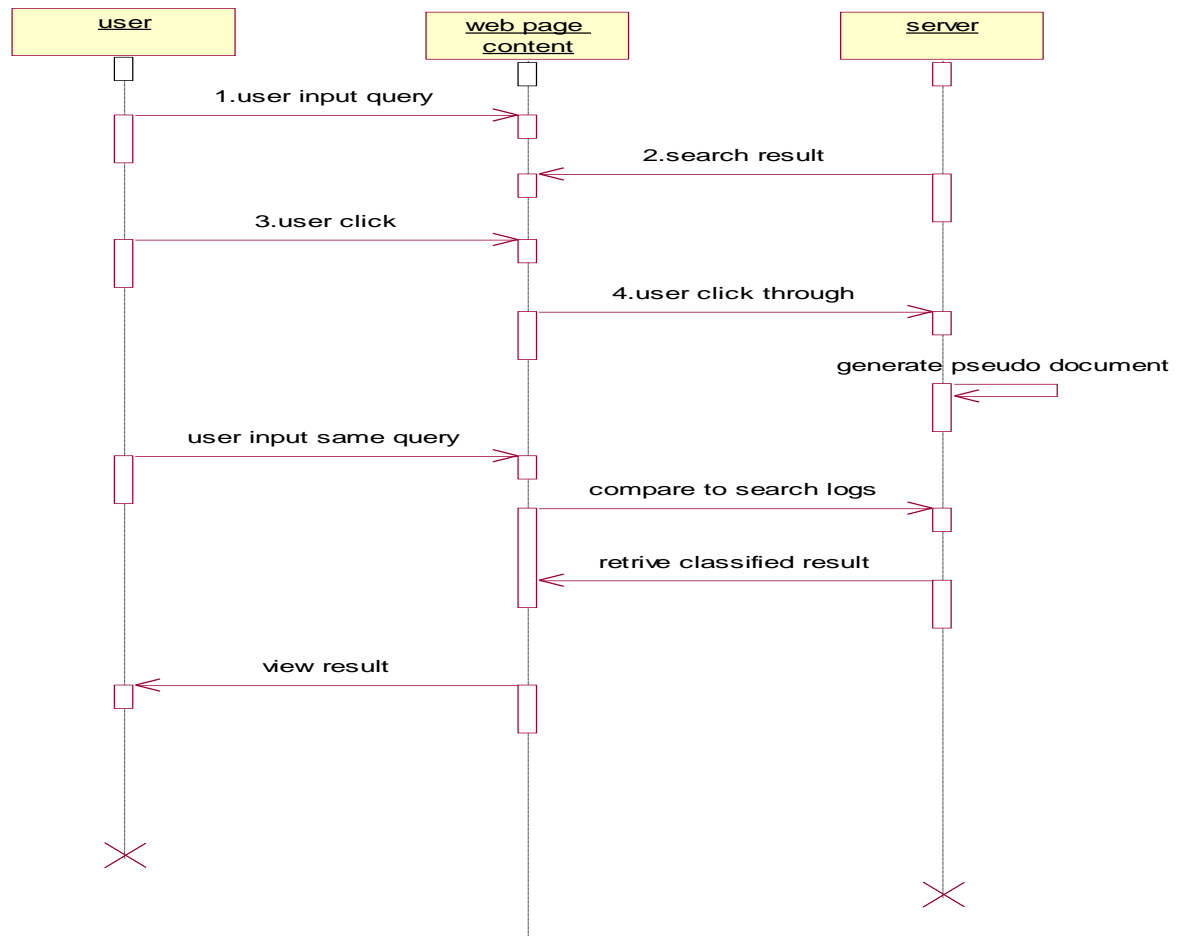


Figure 3.11. Sequence Diagram

3.14. COLLABORATION DIAGRAM

A collaboration diagram, also known as a communication diagram, is a type of UML (Unified Modelling Language) diagram that shows the interactions and communications between objects or components in a system (as shown in Figure 3.12.). It provides a visual representation of how objects or components collaborate to achieve a specific behaviour or functionality.

In a collaboration diagram, objects or components are represented as rectangles, with their names written inside the rectangles. Lines connecting the objects or components represent the communications or interactions between

them. These lines are typically labelled with the type of message being exchanged, such as method invocations, responses, or signals. The messages flow between the objects or components, indicating the order and timing of the interactions.

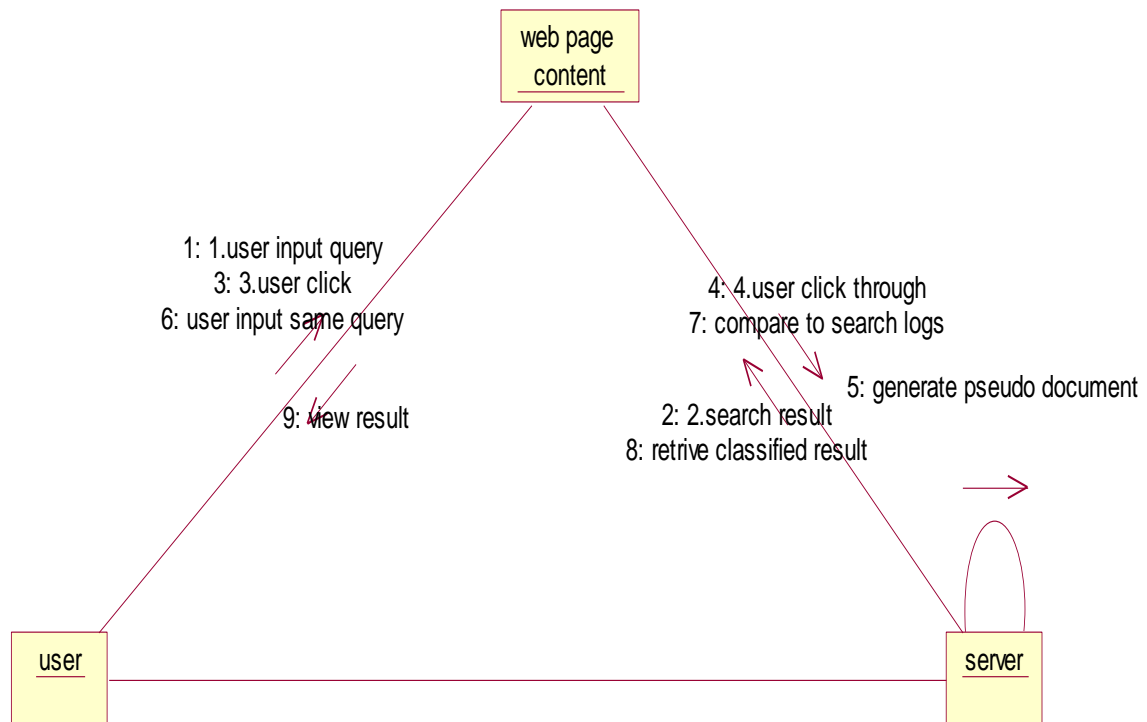


Figure 3.12. Collaboration Diagram

3.15. ACTIVITY DIAGRAM

An activity diagram is a type of UML (Unified Modelling Language) diagram that models the flow of activities or behaviours in a system, such as a business process, workflow, or use case (as shown in Figure 3.13.). It provides a visual representation of how activities are sequenced, coordinated, and executed within a system.

In an activity diagram, activities are represented as rounded rectangles, with the names of the activities written inside the rectangles. Arrows connecting

the activities represent the flow of control between them, indicating the order in which activities are executed. Decision points, represented as diamond-shaped symbols, are used to make decisions and branch the flow of control based on conditions or constraints. Forks and joins, represented as horizontal bars, are used to represent parallel or concurrent flows of control. Swim lanes or partitions can be used to represent different actors or entities responsible for executing the activities.

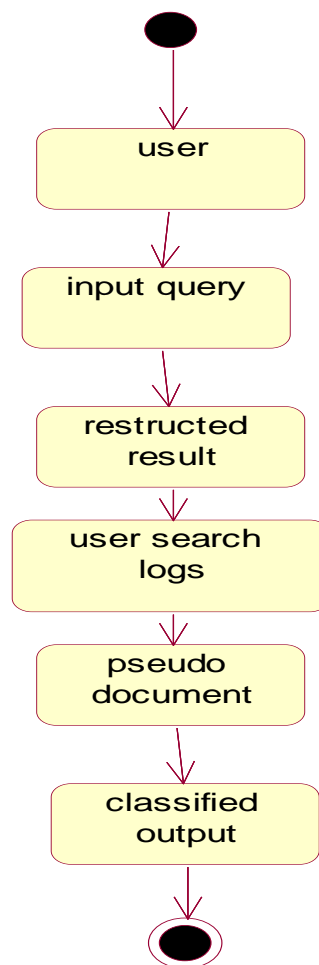


Figure 3.13. Activity Diagram

CHAPTER 4

IMPLEMENTATION AND TESTING

The proposed work is implemented in the following hardware and software setup.

4.1. H/W SYSTEM CONFIGURATION

- **Processor:** Intel Processor
- **Ram:** 4 gigabytes (GB)
- **Hard Disk Space:** 8 gigabytes (GB)

4.2. S/W SYSTEM CONFIGURATION

- **Operating system:** Windows 10
- **Coding Language:** JAVA for backend, JSP for frontend
- **Database:** MySQL
- **Browser:** Firefox (or) Chrome

4.3. SOFTWARE ENVIRONMENT

4.3.1. Java Programming Language

Java is commonly used in the back end of web applications. Java provides a number of tools and frameworks for building server-side web applications, including Java Servlets, JavaServer Pages (JSP), and JavaServer Faces (JSF). These tools allow developers to handle requests from web browsers, process data, and generate dynamic content, which is then sent back to the client-side for display.

Java provides a robust and scalable solution for building the backend of web applications. Java's strong object-oriented programming (OOP) features,

extensive libraries, and community support make it a popular choice for building server-side applications that can handle large amounts of concurrent requests and deliver reliable performance.

4.3.2. JavaServer Pages

JavaServer Pages (JSP) is a Java technology that allows software developers to dynamically generate HTML, XML or other types of documents in response to a Web client request. The technology allows Java code and certain pre-defined actions to be embedded into static content.

The JSP syntax adds additional XML-like tags, called JSP actions, to be used to invoke built-in functionality. Additionally, the technology allows for the creation of JSP tag libraries that act as extensions to the standard HTML or XML tags. Tag libraries provide a platform independent way of extending the capabilities of a Web server. JSPs are compiled into Java Servlets by a JSP compiler. A JSP compiler may generate a servlet in Java code that is then compiled by the Java compiler, or it may generate byte code for the servlet directly. JSPs can also be interpreted on-the-fly reducing the time taken to reload changes. Java Server Pages (JSP) technology provides a simplified, fast way to create dynamic web content. JSP technology enables rapid development of web-based applications that are server- and platform-independent.

4.3.3. Servlets

Servlets are Java classes that run on a web server and handle incoming HTTP requests and generate HTTP responses. They are used in conjunction with JavaServer Pages (JSP) to create dynamic web pages. Servlets are platform-independent and thread-safe, and they are faster and more efficient than traditional CGI scripts. Servlets are widely used in building web applications, especially those that require dynamic content and data access.

4.3.4. MySQL

MySQL uses SQL (Structured Query Language) as its main interface for managing databases, and it supports a range of SQL commands for creating, modifying, and querying databases. MySQL allows for the creation of multiple databases, tables, and relationships between tables, making it a highly flexible and customizable database management system.

MySQL is widely used in web applications for storing and retrieving data, and it is also used in many other types of software systems such as content management systems, e-commerce platforms, and data warehousing systems. It is highly scalable and can handle large amounts of data, making it a popular choice for many organizations.

4.3.5. Stemming Algorithm

Stemming is a process of reducing a word to its base or root form by removing any affixes such as prefixes or suffixes. Stemming algorithms are used to perform this task, and they are an essential part of natural language understanding and natural language processing. Stemming can help extract meaningful information from vast sources of data such as the Internet, and it is commonly used in search engines and information retrieval systems. By reducing words to their base form, stemming algorithms can match different variations of the same word, improving text search accuracy.

4.4. IMPLEMENTATION

Registration Page

The registration page is an essential component of an application that requires user authentication. The page typically includes a form that prompts the user to provide their basic information, such as their username, email, and password (as shown in Figure 4.1.). The purpose of the registration page is to create a unique account for the user so that they can access the full functionality of the application.

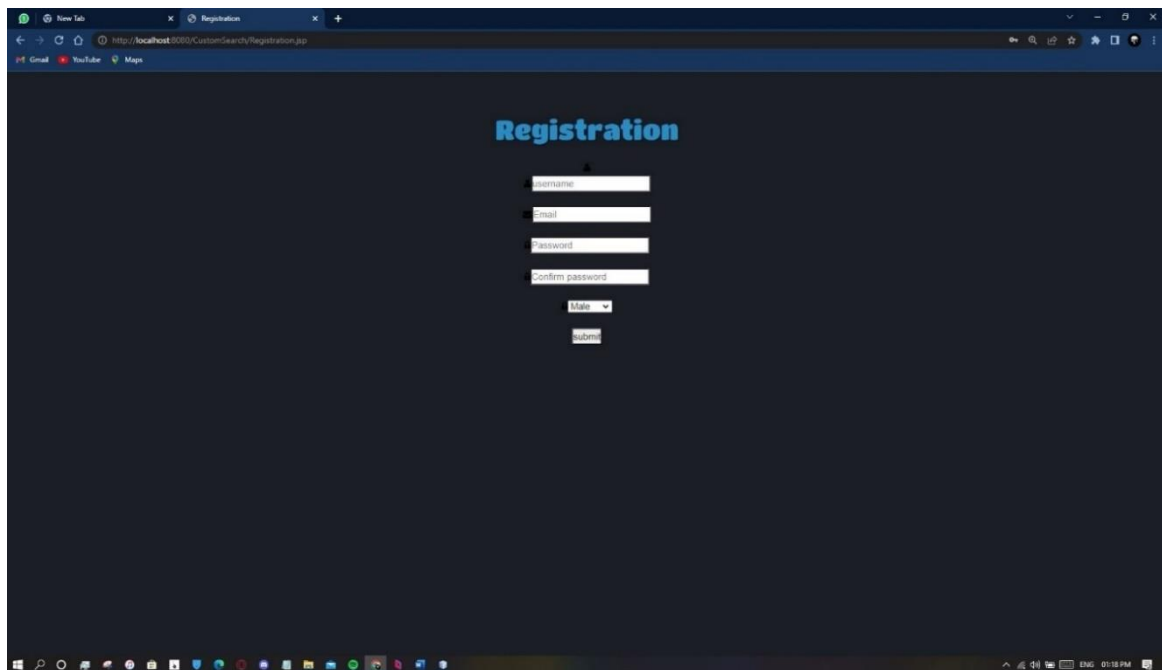
A screenshot of a web browser displaying a registration page. The browser's address bar shows the URL 'http://localhost:5000/CustomSearch/Registration.jsp'. The page has a dark blue background. At the top center, the word 'Registration' is written in a light blue, bold font. Below the title, there is a registration form with the following fields: 'username' (text input), 'email' (text input), 'password' (password input), 'confirm password' (password input), a 'Male' dropdown menu, and a 'submit' button. The browser's taskbar at the bottom shows various application icons and the system clock indicating 01:18 PM.

Figure 4.1. Registration Page

Login Page

A login page is a common feature of an application that requires user authentication. It is typically used to allow users who have already registered to log in and access their account. The login page usually consists of two main fields: the username or email and the password (as shown in Figure 4.2.). These fields are used to authenticate the user and verify their identity.

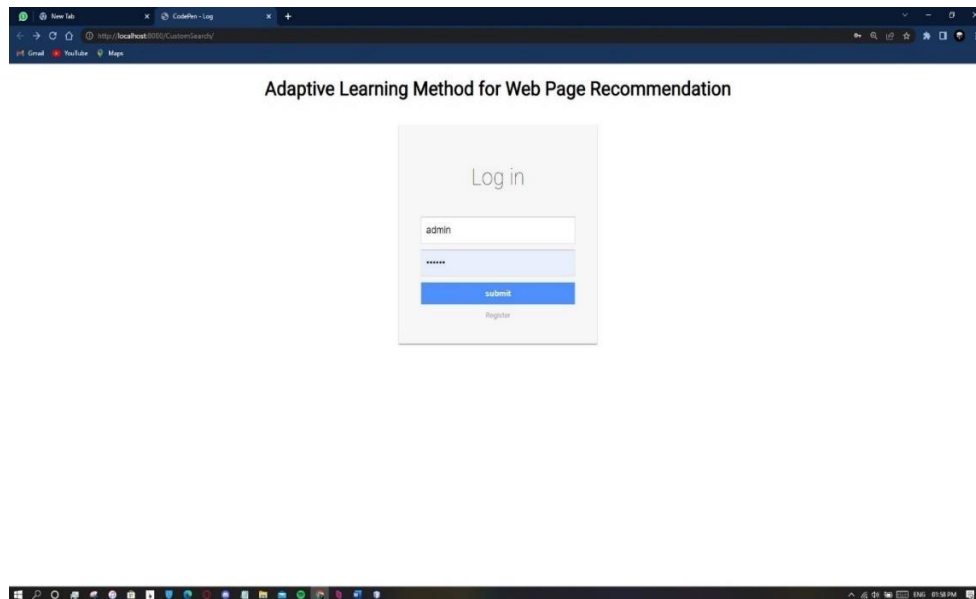


Figure 4.2. Login Page

Search Page

The search box in a search page is where users can enter their search queries or keywords. Once the user enters their query and hits the "search" button, the search engine will process the query and display a list of relevant results on the result page (as shown in Figure 4.3.). The search box is an essential feature of a search engine, as it allows users to quickly and easily find information on the internet.

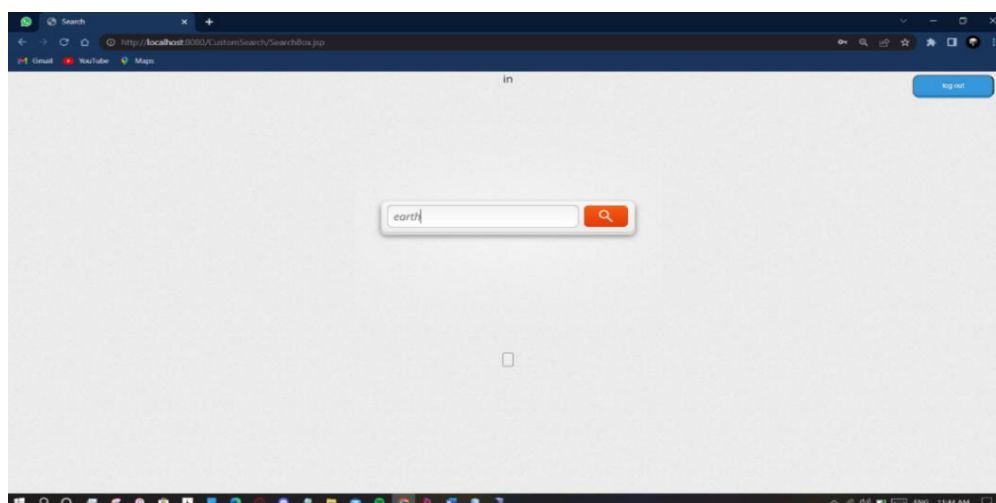


Figure 4.3. Search Page

Result Page

A result page in our application that displays the list of search results after a user enters search terms into the search box and clicks "search" (as shown in Figure 4.4.).

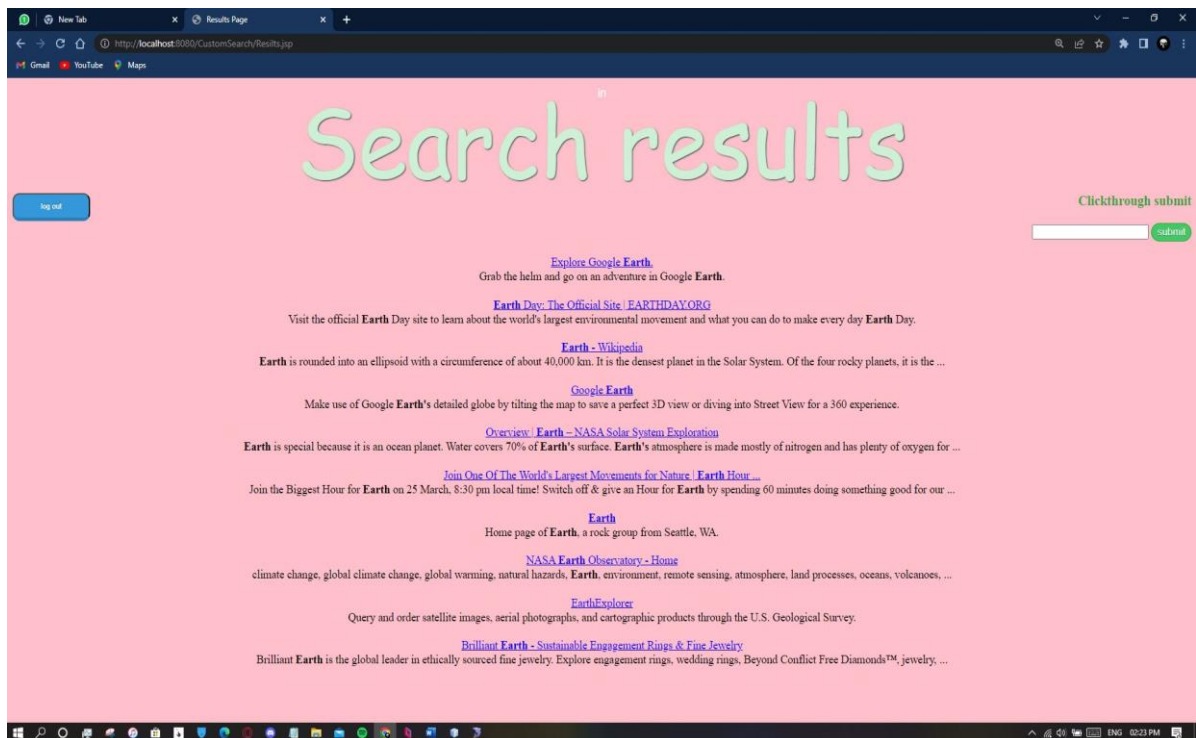


Figure 4.4. Result Page

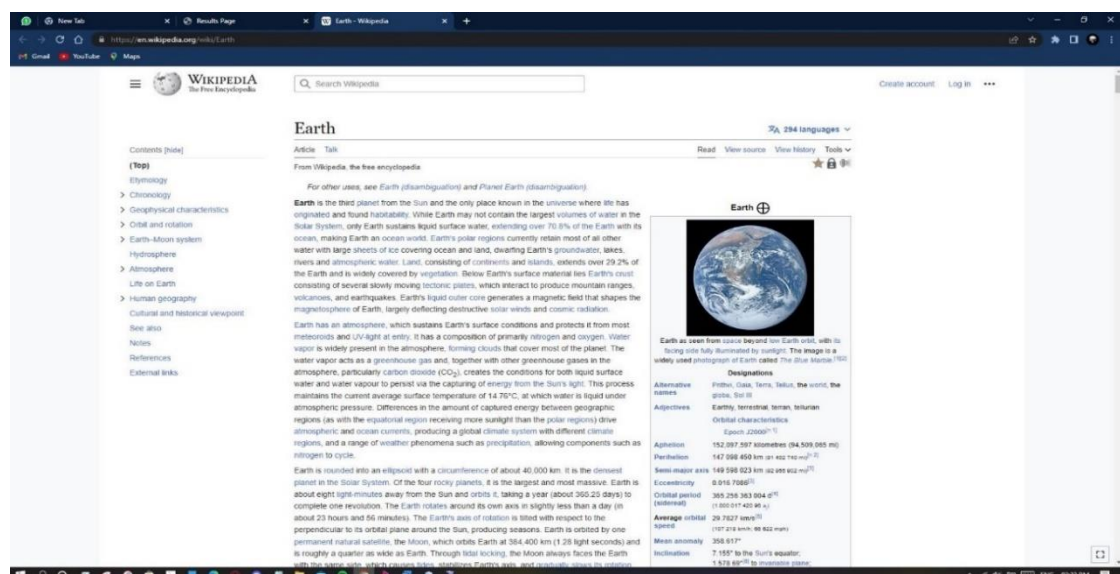


Figure 4.5. Selected Link

Clickthrough Submit

When a user clicks on a specific link in the search engine result page, the corresponding website will be displayed, and the selected link will be highlighted or presented in the clickthrough submission text field (as shown in Figure 4.6.). The user can then decide whether to provide feedback on the relevance and quality of the link's content by selecting the "submit" button. This feedback can help the search engine algorithm to improve the accuracy and relevance of future search results for similar queries.



Figure 4.6. Clickthrough Submit

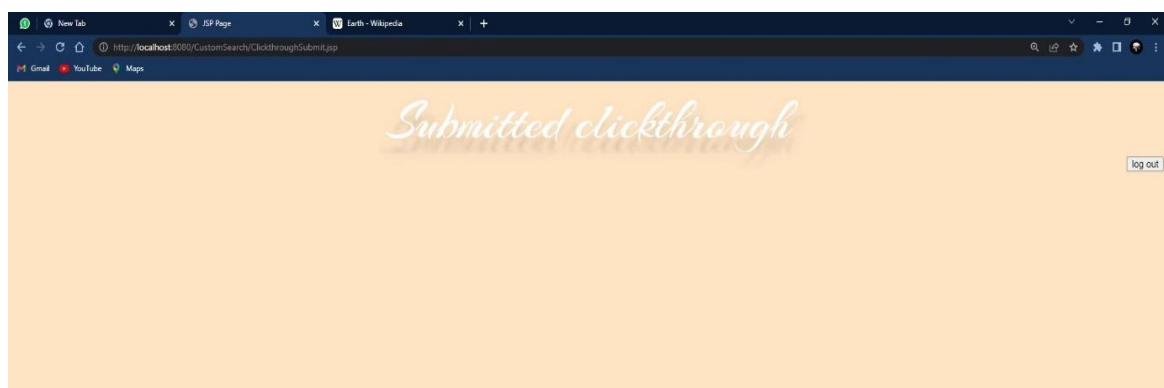


Figure 4.7. Submitted Clickthrough

Recommended Search Result

When a user submits a link in the click-through submission field and then searches for the same query again, the submitted link will appear as the first result in the result page. In addition, the result page will display other links that are relevant to the submitted link, providing the user with a more comprehensive list of resources related to their query (as shown in Figure 4.8.).

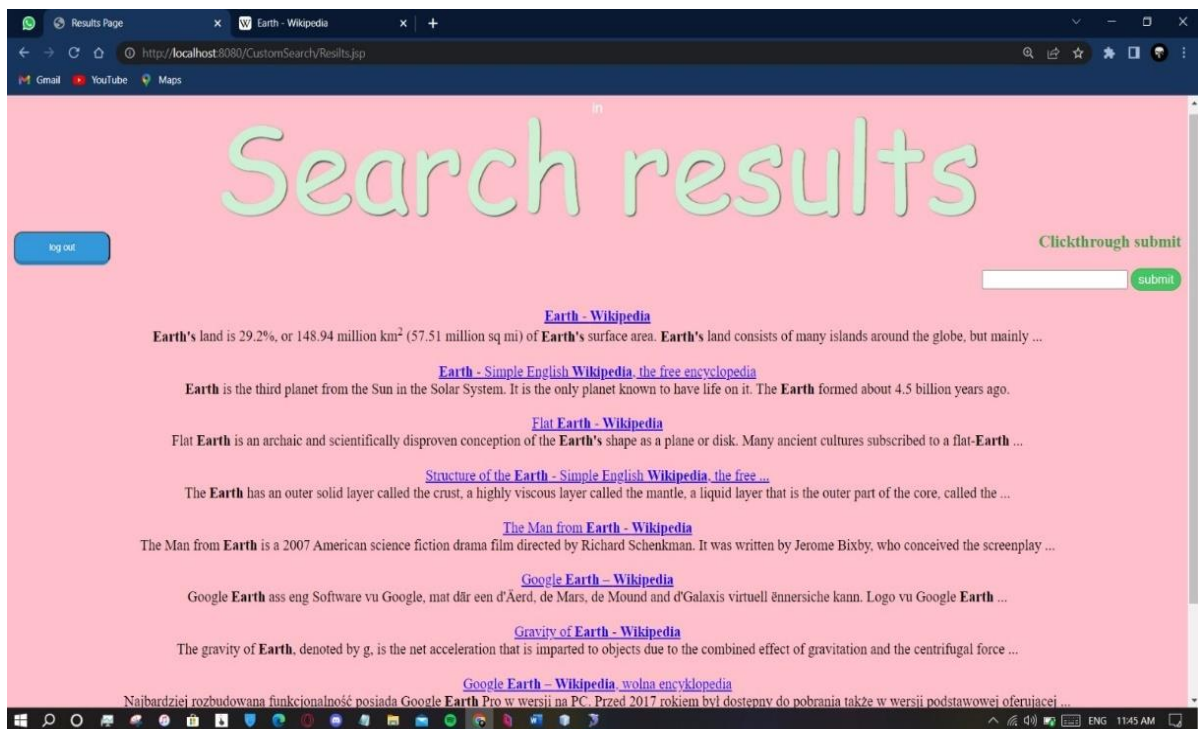


Figure 4.8. Recommended Search Result

4.5. SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product it is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of testing. Each test type addresses a specific testing requirement.

4.6. TYPES OF TESTING

Functional Testing

Functional testing is an important testing technique that can be applied to a wide range of software applications, including search engines. By evaluating the search engine's ability to perform specific functions and meet functional requirements, functional testing can help to identify any defects or issues that may affect the performance or usability of the search engine. Overall, functional testing plays a critical role in ensuring the quality and reliability of search engines and other software applications.

Performance Testing

Performance testing is an important type of software testing that is focused on evaluating the performance and scalability of a system or application. The main goal of performance testing is to identify any potential bottlenecks or issues that may affect the system's performance under various loads and conditions.

In the context of search engines, performance testing is used to ensure that the search engine can handle a high volume of user traffic and perform well under different loads. This involves testing the search engine's speed, responsiveness, and scalability under various traffic levels and workloads.

Performance testing can also involve testing the search engine's ability to handle peak traffic periods and sudden spikes in user activity, as well as its ability to maintain a consistent level of performance over time. By identifying and addressing any performance-related issues, performance testing helps to ensure that the search engine is capable of delivering a fast and reliable search experience for its users.

Usability Testing

Usability testing is a critical part of software testing that is focused on evaluating the ease of use, learnability, and overall user experience of a product or system. The primary goal of usability testing is to identify any potential usability issues or problems that users may encounter while using the product and provide feedback for improving the product's usability.

In the context of search engines, usability testing is used to ensure that the search engine is easy to use and provides a positive user experience. It involves testing the search engine's user interface, navigation, and accessibility to ensure that users can easily find the information they are looking for and navigate the search results.

Usability testing can also involve testing the search engine's response time, user feedback, and error messages to ensure that the search engine provides helpful feedback to users and minimizes frustration. By identifying and addressing any usability-related issues, usability testing helps to ensure that the search engine is user-friendly and provides an excellent user experience for its users.

Relevance Testing

Relevance testing is a type of testing that is used to ensure that the search engine is providing accurate and relevant results to users based on their search queries. It involves testing the search engine's algorithms and ranking methods to ensure that they are working effectively.

In the context of search engines, relevance testing is critical to ensure that the search engine's results are accurate, up-to-date, and relevant to the user's search query. It involves testing various search scenarios and evaluating the search engine's ability to match relevant results to the user's search query.

Relevance testing can also involve testing the search engine's ability to handle synonyms and variations of search terms, as well as its ability to provide context-based results based on the user's location, search history, and other relevant factors. By identifying and addressing any relevance-related issues, relevance testing helps to ensure that the search engine provides accurate and relevant results to its users.

4.7. TEST CASES

Sl. No	Input	Output	Pass/Fail	Expected Result
1	Input password validation Must have at least 6 characters	TRUE (0.568)	Passed	True
2	Login Password Matching	TRUE (0.59)	Passed	True
3	Search Queries	FALSE (0.30)	Passed	False
4	Click Through Submit	TRUE (0.583)	Passed	True

Table 4.1 Test Cases

4.8. RESULTS AND DISCUSSION

4.8.1. Result 1

If the user leaves the search box empty and attempts to initiate a search, a message will be displayed saying 'Please fill out this field' (as shown in Figure 4.9.).

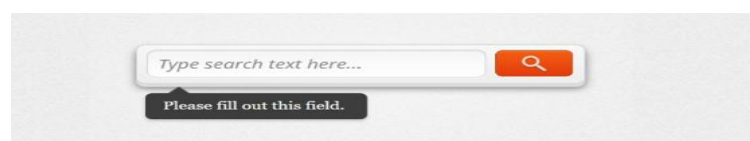


Figure 4.9 Result – False

4.8.2. Result 2

When the user clicks the submit button on a click-through field, the submitted click-through page will be displayed (as shown in Figure 4.10. and Figure 4.11.).



Figure 4.10. Clickthrough Submit



Figure 4.11. Result-True

CHAPTER 5

CONCLUSION

In this work, a novel approach is proposed to infer user search goals for a query by clustering its feedback sessions represented by pseudo-documents. First, feedback sessions are introduced to be analyzed to infer user search goals rather than search results or clicked URLs. Both the clicked URLs and the unclicked ones before the last click are considered as user implicit feedbacks and taken into account to construct feedback sessions. Therefore, feedback sessions can reflect user information needs more efficiently. Second, feedback sessions are mapped to pseudo-documents to approximate goal texts in user minds. The pseudo-documents can enrich the URLs with additional textual contents including the titles and snippets. Based on these pseudo-documents, user search goals can then be discovered and depicted with some keywords. Finally, a new criterion CAP is formulated to evaluate the performance of user search goal inference. Experimental results on user click-through logs from a commercial search engine demonstrate the effectiveness of the proposed methods.

The complexity of this approach is low and can be used in reality easily. For each query, the running time depends on the number of feedback sessions, and therefore, it is usually short. In reality, this approach can discover user search goals for some popular queries offline at first. Then, when users submit one of the queries, the search engine can return the results that are categorized into different groups according to user search goals online. Thus, users can find what they want conveniently.

REFERENCES

1. G.-J. Hwang, H.-C. Chu, and C. Yin, “Objectives, methodologies and Research and issues of learning analytics,” 2017.
2. G. Balakrishnan and D. Coetzee, “Predicting student retention in massive open online courses using hidden Markov models,” *Electrical Engineering and Computer Sciences University of California at Berkeley*, vol. 53, pp. 57–58, 2013.
3. A. Shimada, F. Okubo, and O. Hiroaki, “Browsing-pattern mining from book logs with non-negative matrix factorization,” in *Proc. 9th International Conference on Educational Data Mining*, 2016, pp. 636–637.
4. R. F. Kizilcec, M. Perez-Sanagustín, and J. J. Maldonado, “Self regulated learning strategies predict learner behaviour and goal attainment in massive open online courses,” *Computers & Education*, vol. 104, pp. 18–33, 2017.
5. G. Akc,apınar, M. N. Hasnine, R. Majumdar, B. Flanagan, and H. Ogata, “Developing an early-warning system for spotting at-risk students by using eBook interaction logs,” *Smart Learning Environments*, vol. 6, no. 1, pp. 1–15, 2019.
6. H. Drachsler, K. Verbert, O. C. Santos, and N. Manouselis, “Panorama of recommender systems to support learning,” in *Recommender systems handbook*. Springer, 2015, pp. 421–451.

7. K. Thaker, L. Zhang, D. He, and P. Brusilovsky, "Recommending remedial readings using student knowledge state," in Proc. 13th International Conference on Educational Data Mining, 2020, pp. 233–244.
8. T. Nakata, "Does repeated practice make perfect? the effects of within session repeated retrieval on second language vocabulary learning," *Studies in Second Language Acquisition*, vol. 39, no. 4, pp. 653–679, 2017.
9. R. V. Lindsey, J. D. Shroyer, H. Pashler, and M. C. Mozer, "Improving Students' Long-Term Knowledge Retention Through Personalized Review," *Psychological Science*, vol. 25, no. 3, pp. 639–647, 2014.
10. T. Shiino, A. Shimada, T. Minematsu, and R.-i. Taniguchi "Learning support through personalized review material recommendations," in Proc. 28th International Conference on Computers in Education, 2020, pp. 137–143.

PUBLICATION:

Gokul Sri Ram S, Narash Kumar A, Edwin Rajan G, Bala Vignesh G & Dr. J. Benadict Raja (April 2023). "Adaptive Learning Method For Web Recommendation". *International Research Journal of Modernization in Engineering Technology and Science*. (Volume 5 Issue 4, Sr.No:285). <https://www.doi.org/10.56726/IRJMETS35758F>