

Addis Ababa
University
(Since 1950)



**ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
SCHOOL OF INFORMATION SCIENCE**

**APPLICATION OF CASE BASED RECOMMENDER SYSTEM
TO ADVISE STUDENTS IN FIELD OF STUDY SELECTION AT
HIGHER EDUCATION IN ETHIOPIA**

By

Biazen Getnet

January, 2013

**ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
SCHOOL OF INFORMATION SCIENCE**

**APPLICATION OF CASE BASED RECOMMENDER SYSTEM
TO ADVISE STUDENTS IN FIELD OF STUDY SELECTION AT
HIGHER EDUCATION IN ETHIOPIA**

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR DEGREE OF MASTER OF SCIENCE IN INFORMATION SCIENCE

BY
Biazen Getnet

**ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
SCHOOL OF INFORMATION SCIENCE**

**APPLICATION OF CASE BASED RECOMMENDER SYSTEM
TO ADVISE STUDENTS IN FIELD OF STUDY SELECTION AT
HIGHER EDUCATION IN ETHIOPIA**

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR DEGREE OF MASTER OF SCIENCE IN INFORMATION SCIENCE

By

Biazen Getnet

Signature of board members for approval

Advisor -----signature-----

Chairperson -----signature-----

Examiner -----signature-----

DECLARATION

This thesis is my original work and has not been submitted as a partial requirement for a degree in any university.

Biazen Getnet
January, 2013

The thesis has been submitted for examination with our approval as university advisors.

Gashaw Kebede (PhD)
January, 2013

Acknowledgement

First of all, I would like to forward my heartfelt gratitude to my research advisor, Dr. Gashaw Kebede, for his crucial guidance and encouragement since the conception of this research work.

I would like to thank many people for their kindness and assistance during my thesis work. I truly appreciate the support of the people at the National Agency for Examination and Debre Markos University.

I wish to thank my beloved friends and classmates who have one way or another contributed to the success of my thesis.

Special thanks go to Ato Haleluya Kefilu whose brilliance inspires me in all that I do. I owe a lot for his patience, encouragement and heart-warming personality.

My heartfelt thanks also go to my instructor, Dr. Million Meshesha, for his lovely time and classes and his special encouragement for me to do my research in this area.

Above all, I thank God who surrounded me with so many wonderful people in all my life.

Table of Contents

LIST OF FIGURES.....	V
LIST OF TABLES.....	VI
LIST OF ACRONYMS.....	VII
ABSTRACT.....	VIII
CHAPTER ONE	1
INTRODUCTION.....	1
1.1. Background of the Study.....	1
1.2. Statement of the Problem and Justification	4
1.3. Objective of the Study.....	7
1.3.1. General Objective	7
1.3.2. Specific Objectives	7
1.4. Research Methodology	8
1.4.1. Literature Review.....	8
1.4.2. Data Collection Methods	8
1.4.3. Programming Technique and Tools	10
1.4.4. Evaluation/Testing procedures.....	11
1.5. Scope and Limitation of the Study	11
1.6. Significance and Application of the Study.....	12
1.7. Organization of the Thesis	13
CHAPTER TWO	14
REVIEW OF RELATED LITERATURE	14
2.1. Definition of Recommender System	16
2.2. Architecture for Recommender System	17

2.3.	Techniques applied in Recommender System	19
2.3.1.	Content Based Recommender System	20
2.3.2.	Collaborative Filtering Recommender System	23
2.3.3.	Knowledge based recommender systems	27
2.3.4.	Hybrid Recommender System	28
2.4.	Process of Recommender System.....	29
2.5.	Input of Recommender System	30
2.6.	Output of a Recommender System.....	31
2.7.	Case Based Recommendation.....	31
2.8.	Knowledge Representation	32
2.9.	Student placement process.....	35
2.9.1.	Student`s Placement Process in Ethiopia	35
2.10.	Related works	37
CHAPTER THREE		41
KNOWLEDGE ACQUISITION AND MODELING		41
3.1.	Overview Knowledge Base System	41
3.2.	Architecture of a Knowledge Base System	42
3.3.	Knowledge Base System Development.....	43
3.3.1.	Knowledge Acquisition.....	44
3.3.2.	Knowledge Modeling	47
3.4.	Hierarchical structure of field of study	48
CHAPTER FOUR		55
DESIGNING AND IMPLEMENTATION OF THE PROTOTYPE		55
4.1.	Designing the Architecture of Case Based Recommender System.....	56

4.2. Case Based Recommender System for field of study selection	58
4.2.1. Building the Case Base	58
4.2.2. Case Representation	59
4.2.3. Managing Case structure	64
4.2.4. Managing Connectors	65
4.2.5. Managing Tasks/ Methods.....	66
4.2.6. Deploy the case base recommender system	71
CHAPTER FIVE	73
TESTING AND PERFORMANCE EVALUATION OF THE PROTOTYPE	73
5.1. Experimental Setting	73
5.2. Testing the main Cycles and Evaluating the Performance of the System	74
5.2.1. Evaluation of the Retrieval and Reuse Process	74
5.2.2. Case Revision and Solution Adaptation Testing	77
5.2.3. Case Retaining testing.....	78
5.2.4. User Acceptance Testing.....	79
CHAPTER SIX.....	84
CONCLUSION AND RECOMMENDATION	84
6.1. Conclusion	84
6.2. Recommendation	87
REFERENCES	89

LIST OF FIGURES

Figure 2. 1 User interactions on the web (source: Ning , 2007).	17
Figure 2. 2 Architecture of recommender system (source: Terveen and Hill, 2001)	18
Figure 2. 3 Recommender System Process (source: Schafer et al, 2001)	19
Figure 2. 4 Content based recommender system (source: Semeraro, 2010).....	21
Figure 2. 5 Collaborative Filtering Recommender System. (Source: Asanov, 2011).	24
Figure 2. 6 Collaborative Filtering Recommender System (source: Mobasher,2002)	25
Figure 4. 1 Case Based Recommender System Architecture.....	57
Figure 4. 2 JCOLIBRI main windows	58
Figure 4. 3 Managing Case Structure	65
Figure 4. 4 JCOLIBRI case base schemas	65
Figure 4. 5 Managing Connectors	66
Figure 4. 6 case similarities between case base and query.....	68
Figure 4. 7 Revision and retain tasks	69
Figure 4. 8 Managing Tasks/Methods.....	70
Figure 4. 9 Case entry windows	71
Figure 5. 1 Retrieval case	75
Figure 5. 2 case revision testing.....	78
Figure 5. 3 Case Retaining windows	79

LIST OF TABLES

Table 3. 1 Field of studies under Engineering and technology.....	50
Table 4. 1 Selected attributes by domain experts	60
Table 4. 2 Descriptions and Weight of the Selected Attributes	63
Table 5. 1 Relevant Cases Assigned by the Domain Expert for Sample Test Cases.....	76
Table 5. 2 Performance Measurement of case base recommender system using Precision and Recall.....	77
Table 5. 3 performance evaluation by domain expert	80
Table 5. 4 performance evaluation of the system by students	82

LIST OF ACRONYMS

AI: Artificial intelligence

CBR : Case-Based Reasoning.

GAIA: Group of Artificial Intelligence Applications

IEEE: Institute of Electrical and Electronics Engineers

JCOLIBRI: java Cases and Ontology Libraries Integration for Building Reasoning Infrastructure

SWI-Prolog: *Sociaal- Wetenschappelijke Informatica* (Social Science Informatics)
programming in logic.

TV: television

ABSTRACT

Using recommender systems with the help of computer systems technology to support the academic advising process offers many advantages over the traditional student advising system. One of the main problems faced by students is to take the right decision in relation to field of study selection process based on available information. The objective of this research is to develop a prototype case base recommender system that assists the students in their field of study selection process. The system provides recommendation to the students based on previously solved cases and new query given by the student.

For this study, about 105 cases which are collected from successful students and 13 attributes which are collected from experts are used as case base. These attributes and cases are used as knowledge base to construct case base recommender. The system calculates similarity between existing case and new queries that are provided by the students and provides solution or recommendation by taking best cases to the new query. This recommendation enables students to make decision easily. In this study, JCOLIBRI case base development tool is used to develop the prototype of case based recommender system. JCOLIBRI contains both user interface which enables students to enter their query and programming codes with the help of Java script language.

After developing the prototype of the system, testing of the prototype for case base recommender system was done to evaluate the performance of the system. Based on prototype testing, the average performance of the system is 77.2% and 80.2% by the domain experts and students respectively.

Using field of study selection recommender system for students helps them to make decision easily to select appropriate field of study based on their interest and background information. The researcher recommends for the future work to improve the system that can recommend university as well as field of study selection by providing detail information about each selection.

CHAPTER ONE

INTRODUCTION

1.1. Background of the Study

We are often making choices without sufficient personal experience of the alternatives that are available to us in different circumstance. In our everyday life, we sometimes rely on recommendations from other people either by word of mouth, recommendation letters or on movie and book reviews to select from the huge amount of information that is available in different places. But these suggestions are not enough in this digital age (Symeonidis et al., 2007). Currently, Recommender systems could be found in many modern applications that expose the user to huge collections of items both commercially and in the research community, where many approaches have been suggested for providing recommendations. Such systems typically provide the user with a list of recommended items they might prefer, or predict how much they might prefer each item. It helps to match the users with items to ease information overload; making decision about items and sales assistance (Guy and Asela, 2010).

Recommender systems are a type of information filtering system that gives advice on products, information, or services that a user may be interested in. They assist users with the decision making process when choosing items with multiple alternatives (Werthner, Hansen and Ricci 2007).

Recommendation system can be classified into traditional and web based recommender system. Traditional (manual) recommendation system such as peer and family advice suffers from the accurate perdition of matching the right decision.

Service, product, information and item, suggestions provided by recommender systems are useful when users are overwhelmed by a large number of options to consider or when they do not have enough knowledge about a specific domain to make autonomous decision (Adomavicius and Tuzhilin, 2005).

Recommendation process is a sequential process. At each point the recommender system makes a decision about the item/items. This decision should take into account the sequential process involved and the optimization criteria suitable for the recommender system (Zimdars et al., 2001). They suggest items to the user who can then accept one of the recommendations. At the next stage a new list of recommended items is calculated and presented to the user. This sequential nature of the recommendation process, where at each stage a new list is calculated based on the user's past ratings, will lead us naturally to our reformulation of the recommendation process as a sequential optimization process (Guy et al.,2005).

Recommender systems may be based on collaborative filtering (by user ratings), content-based filtering (by keywords), knowledge based recommender system that uses knowledge about users and products to pursue a knowledge based approach to generating a recommendation by reasoning about what products meet the user's requirements and hybrid filtering (by both collaborative and content-based filtering) (Burke, 2006, Schafer et al, 2001).

Case based recommender system is a part of knowledge based recommender system that depends on case based reasoning to generate personalized recommendations for exploiting the knowledge contained in past recommendation cases. These systems assume that the quality of a new recommendation depends on the quality of the recorded recommendation cases (Aamodt, 1994, Burke, 2007). A case based recommender system maintains a set of cases of previously solved recommendation problems and their solutions. According to Shimazu (2002), case base is the product catalogue which is the solutions of the recommendation problem, and the problem is the user's query that is essentially a partial description of the desired product. A case could be thought of as a record in a database; collection of features and their values.

The effectiveness of the recommendation in case base recommender system is based on the ability to match user preferences; with product description; the tools used to explain the match and to enforce the validity of the suggestion; the range of available functionalities and the graphical interface that support the user in browsing the information content, either the cases or the products to recommend (Fabiana and Francesco, 2003).

A university education has become a basic part of most people's preparation for working life. Admission to university is therefore a topic of importance. How a student chooses a university and field of study and also conversely how a university chooses a student, determines the success of both sides in carrying out the education and to have skilled man power for the country.

In governmental higher education institution in Ethiopia, as some percent of educational and living expenses are covered by the government, placement is done by Ministry of Education based on the score of the student, the availability of fields of study, the number of students that it can accommodate, intake capacity of the universities, the government policy, proximity of the institution to student's home etc (Getachew, 2008).

However, most existing admission processes are based on the perspective of universities to receive the new incoming students and student's interest based on their results but not on the perspective of preparatory schools that are sending their students to pursue higher education. The university knows very little about the applicant, the secondary school knows a great deal more about their student. After admission, students are expected to have good performance in the studies and graduate with good knowledge to solve the problem of the country with their domain knowledge acquired from the university (Luan, 2002).

There are factors that determine placement of students to the university and field of study they are interested in (Ebrahim, 1999). Some of them include the following.

- Available intake capacity of universities
- Minimum educational requirement (result) to assign their field of study and university.

According to Ministry of Education (National Agency for Examination of Ethiopia), the placement of students into field of study and university is based on their higher education entrance examination results. Those scoring higher results than others can have the chance to be assigned to university and field of study based on their choice. Students have the responsibility to select their field of interest and university based on their score of higher education entrance examination. As per the revised manual of higher education institution

student's placement, student's have the right to choose any field of study and university in order of applicant's interest (Ministry of Education, 2011). Even though national agency for examination tries to consider students' interest and intake capacity of universities, there are always students who are assigned to field of study which is not their choice or interest. Such students complain about and become disappointed with the field they are assigned to. This condition brings about negative impact on their successes of tertiary level studies.

1.2. Statement of the Problem and Justification

According to Ministry of Education of Ethiopia, student who have completed their preparatory class and fulfilled minimum result requirement of the year will choose their field of study and university based on their interest to join their higher education. However, a major problem student face is selecting a field of study that is appropriate for them to pursue in higher institutions (Getachew, 2008).

Students are required to select their field of study to join higher institution after completing their preparatory class. This selection process can be grouped into two (Ministry of Education, 2011): band selection and department selection. In the first process, students are required to complete band or sub- band selection in order to join different schools or faculties in the different higher education institutions in the country. National Agency for Examination grouped field of studies in the form of bands and sub-bands. For example, band1 represents engineering and technology, sub-band1A represents engineering science, sub-band1B represents computer science etc.

The second process is selecting departments after joining university. This process is distributing students to different departments after they joined field of study or bands and sub-bands. The main objective of this study is to address the problems that students face in selecting bands and sub bands as this is the stage where students face the first problem in the process of selecting appropriate field of study.

According to the interview made with the director of yekatit 12 preparatory school, preparatory schools are overloaded with students who made wrong choices and need re-application form. The director replied that there are many students that ask the school every year for re-application due to misunderstanding and wrong selection of bands and sub-bands they made.

According to the interview, most of the students completing the preparatory level do not know the detail of these bands and sub-bands and they need advice to select field of study based on their interest. School directors also do not have adequate knowledge about the bands and sub-bands to advice students when they select field of study. This leaves most of the students confused to know about each field of study and make the selection blindly.

Further interview with students of Debre Markos University engineering (former mathematics) and Addis Ababa University Information Systems indicated that most of the students choose their field of study blindly without compressive advice from expert or recommender system. To some extent they follow advice from families, peers and their senior and graduated students from universities. However the advice from these different groups did not help them much in selecting a field of study that best fit their background knowledge and skills.

To make things even worse, in the rural part of the country, it is known that students are unable to get relevant information and advice about how to choose field of study and they usually select blindly. Based on the interview made with head of psychology department at Debre Markos University, most students came from rural part of the country are affected with this problem since they select their field of study without relevant information.

An evidence for the existence of problems in field of study selection, one can observe the increasing number of posts about exchange of university and field of study made by dissatisfied students on different announcement boards and communication Media. For instance, in 2011, after announcement of placement result by the agency; we have seen many posts for field exchange around the surrounding of Addis Ababa University and some preparatory schools. This indicates that there is a problem related to field of study selection. This is due to partly wrong choice of field of study made by students without appropriate advice from expert

system. If students dissatisfied with their assigned field of study are unable to change, they will not be interested to give due attention to their classes. To overcome this problem, students need advice that enables them to decide which field of study is appropriate for them based on their background interest, skill and ability to the specific area of interest.

Recommender systems are considered to address solutions for these problems. For example, recommender systems are being used for course selection by students. As discussed by Ekdahl et al., (2002) students are required to choose the course they are interested in reading for subsequent semesters. It is the student's responsibility to make sure that (s)he gets the course needed to graduate. If the choice is not made even for the mandatory courses, the student will not be allowed to take that course. Similarly, Simon and Robert (2009) tried to design a recommender system that makes different kinds of prediction based on student's information. As per his discussion, University or college admission is a complex decision process that goes beyond simply matching test scores and admission requirements. Almost all admission and enrollment studies are based on the perspective of Universities or colleges, and only few studies are based on the perspective of secondary schools.

In our country Ethiopia, there is no recommender system developed to assist student when they applied for higher education. The only research done is "student's placement and admission expert system" by Ebrahin Kassa (1999). This study was conducted on Addis Ababa university evening students. It did not include regular students of Addis Ababa University that are assigned by Ministry of Education. A closely related research has been done by Getachew Feleke (2008) entitled "Higher Education Entrance Student Placement Processing and Retrieval System" focusing on placement process and how the students can access their placement result online. But it doesn't cover the problem of students choice of field of study.

The aim of this study is therefore to design a case based recommender system that helps students in making informed band and sub-band selection after completion of their preparatory class. This work enables students to solve the problem of field of study selection based on recommendation given by the system.

The study tried to address the following research questions.

- What are the main criteria used by the recommender system to advise students in their field of study selection process?
- How to develop an effective prototype model for field of study recommender system?
- What is the contribution of case base recommender system over manual system for student in the selection of field of study?
- What is the importance of existing cases for the students in their field of study selection process?

1.3. Objective of the Study

The general and specific objectives of the study are the following.

1.3.1. General Objective

The general objective of the study is to develop Case Based Recommender System that advises students in their choice of appropriate field of study or band and sub-band to join higher education institution of Ethiopia.

1.3.2. Specific Objectives

With the above general goal in mind the research tried to fulfill the following specific objectives.

- To review literatures and understand the concept of recommender system and how it is designed.
- To identify and capture the previous cases, facts, insights and rules that students need to know in order to make informed selection of field of study in Ethiopian context.
- To construct a case structure comprising relevant attribute that have a direct impact on students' field of study selection.
- To build the case base comprising different cases in order to implement recommender system model.

- To develop prototype model for student's field of study selection recommender system.
- To evaluate the performance of recommender system.

1.4. Research Methodology

The following methods have been utilized in the process of conducting the study.

1.4.1. Literature Review

Relevant literature such as books, journal articles, conference papers and resources from internet that are related to recommender system especially field of study selection and students' placement system and manuals of development tools and techniques used for system development were reviewed.

In addition to this, documents related to placement of students by Ministry of Education, universities and preparatory schools were also reviewed. Moreover, to assess what others have done in the area and to understand the problem better, related research works and projects were assessed. In general, comprehensive investigations of available literature on different techniques of student placement recommender system were made.

1.4.2. Data Collection Methods

Frequent contacts were made with National Agency for Examination of Ethiopia to assess the existing placement techniques. Data collection methods such as interview and documentary analysis were used, interview with concerned bodies from National Agency for Examination, Debre Markos University and other experts from each area of study including students from selected departments as well as review of documents and manuals including website of the agency have been used to gather relevant information for the development of recommender system. For these data collection methods, non-probability sampling techniques called purposive sampling methods were used for all interviewee groups. This is because only few students and university staff members with the experience relevant to the study had to be selected for the interviews since they are target groups that affected by the system. Similarly from the Ministry of Education, there were target staffs concerned with student's placement

process who were interviewed to get relevant bands and sub-bands related information. To construct relevant cases for the recommender system, selected students and heads of departments from Debre Markos University were interviewed. To identify attributes for the recommender system, interview was made with 32 department heads of Debre Markos University. The aim of the interview with department heads is to find out what they believe on important attributes for students to select field of study that enables them to succeed in their respective bands and sub-bands since each department heads are representative of their respective faculties.

These criteria can be used as the main attributes for the researcher and about thirteen attributes were selected. The researcher prepared interview questions for all 32 department (see appendix F) heads to get relevant attributes. These department heads were selected purposefully to get relevant attributes. From the side of students, interview was made with 105 students selected from 32 departments of Debre Markos University. The selected students are those who are happy with the field of study that they are assigned to and are been successfully performing in their studies. These students were asked to complete a tabular form (see **Appendix A**) that contains columns of attributes identified by the domain experts. The data that each of the 105 students provided against the different attributes is used to create the 105 cases to be used by the recommender system. The researcher used face to face communication in the form of interview question to collect relevant attributes and cases from both experts and students since it needs more explanation by the researcher.

After collecting main attributes and cases, creating knowledge base was the crucial process. The system focused on the side of student's problem in the selection of field of study. Based on the cases collected as an input, recommender system will create knowledge base to provide recommendation for the students to make decision easy for them. The knowledge base contains the following attributes: Preparatory class attend, special skill, interest1, interest2, interest3, background interest, family interest, preparatory best result, entrance exam result, health status, language dependency, duration, past experience and solution attributes.

Cases are represented in the form of case base in text file that are important for make decision by the system. Similarly, there is information given by the students that enable the system to make decision. Finally the system provides recommendation by measuring similarity of existing cases and new query.

1.4.3. Programming Technique and Tools

A knowledge base recommender system depends on either explicit domain knowledge about the items or knowledge about the users to derive relevant recommendation; content based recommender system which builds a profile of keywords from items users like and recommend new items which match the profile and collaborative filtering that finds similar users to a target user by comparing users opinions of items (Bamshad, 2002). In this way, both content based and knowledge based recommender system techniques were used since the recommender system developed in this study depends on domain knowledge of experts that are relevant to the students in their field of study selection process and the contents or profiles used by others. The contents or profiles used by others were used as a case to create new profile which is used as recommendation.

There are many programming tools used to develop recommender system. Among this, SWI-prolog is one of the most powerful declarative programming language in artificial intelligence research and most user friendly tool for knowledge base system development. When implementing the solution to a problem instead of specifying how to achieve a certain goal in a certain situation, we specify what the situation (rules and fact) and the goal (query) and let the prolog interpreter drive the solution for us (Endriss, 2000).

Writing a program in Prolog means writing facts and rules which together comprise knowledge base. SWI-Prolog is license free Prolog compiler (Grzegorz and Nalepa, 2007).

Everything in prolog is defined in terms of two things: the fact and the rule. A fact is a statement consisting simply of an identifier followed by an n-tuple of constants. The identifier is interpreted as the name of a relation and the fact states that the specified n-tuple is in the relation and a relation identifier is referred to as a predicate. A fact states that a certain tuple of

values satisfied a predicate unconditionally. On the other hand, a rule in prolog is a statement which gives conditions under which topples satisfy a prediction (Endriss, 2000).

For this research JCOLIBRI Case base recommender system framework was used since it is most appropriate programming tool for case base systems and based on the following capabilities.

- JCOLIBRI is extensible, reusable, and used by different types of users and for different purposes (development, research and/or teaching), compatible with commercial applications.
- Since it is just a .jar file, it is suitable for web applications.
- More user friendly than others and it works well in external database such as plain texts.
- Suitable for developing large scale applications.

1.4.4. Evaluation/Testing procedures

Domain experts and students were selected to evaluate the developed system. Test cases were selected by the researcher and relevant cases were selected by experts and students to test selected test case by entering their own query. Recall and precision value of the system have been calculated based on its retrieval results. In addition to that, user acceptance testing was performed to test it application by the user and relevant feedbacks were collected to calculate its performance.

1.5. Scope and Limitation of the Study

Student placement can be viewed from two directions; University placement and field of study selections. When students choose their university and field of study, there is also another criterion from the side of National Agency for Examination, namely, the intake capacity of universities and special consideration for disabilities. The proposed recommender system focused on the choice of students' field of study that enables them to choose a field of study that interests them.

Even though there are universities that include different sub-bands and field of studies in different bands, this study focused on universities that are organized based on band structure of Ministry of Education.

The limitation of the study is mainly due to programming tool the researcher used. The system provides only one recommendation for the user. This is because the JCOLIBRI programming tool uses cases to recommend to the user based on the new query and only one recommendation supported with it. In addition to this programming tool problem, there were also problems to get relevant attributes and cases from experts and students respectively. Due to this few departments which are not given in nearby university, the researcher did not include under the case. But when new field of study is required which is not included in the case, the system will store to the case base and used for decision making for the next processing. Getting inadequate information about students to get relevant cases was also another limitation of the study.

1.6. Significance and Application of the Study

From this study, primarily students who join higher education institution, specifically governmental universities and colleges, are the immediate beneficiaries of the developed prototype recommender systems. The prototype also has great significance for students' families, higher institutions to get well interested students, national agency for examination to reduce their placement process time and effort and the country to get interest based graduate.

Field of study selection advice system enables students to choose their field of study based on their interest and background information. The system will give recommendation for students by extracting information from knowledge base system stored as case base. It is also important for National Agency for Examination since it will give full information about the student background and his/her interest of area that enables the agency to assign students based on their interest. The agency also can assign students to their field of study based on the suggestion given by the recommender system.

1.7. Organization of the Thesis

This thesis is organized into six chapters: the first Chapter briefly discusses about the introduction part of the study, background of the problem area, the general and specific objectives of the study, the research methodology, the scope and limitation, programming tools used to develop the prototype, evaluation procedure and significance and application of the results of the research. Chapter two gives an overview of related research work and background introduction about recommender system. Chapter three focuses on the knowledge representation methods and creation of knowledge modeling to develop recommender system. Chapter four focuses on designing and implementation of the prototype which discusses the detail of developed system using selected programming tool. Testing and evaluation of the system by domain area is discussed in the fifth chapter and finally chapter six gives final conclusions of the research done and forward recommendations for future studies.

CHAPTER TWO

REVIEW OF RELATED LITERATURE

Since the invention of internet, many e-resources have been easy to access using the web, like buying or sharing of books (from e-shops), borrowing books from digital library, downloading software, reading journals and articles, doing business using different e-commerce tools. But in recent times because of the exponential growth on the quantity of information, searching for this information on the internet is overwhelming the users; this is called "information overload". And due to the growth of the resources on the web servers of many sites it is hard to display all available data on the website so that the users can access (Adomavicius, 2005).

This age is the age of information abundance. The 1990s have seen as an explosion of information and entertainment technologies, and thus of choices a person faces. People may choose from dozens to hundreds of television channels, thousands of videos, millions of books, CDs, and multimedia, interactive documents on the Web, educational materials and seemingly countless other consumer items presented in catalogs or advertisements in one medium or another. The web in particular offers myriad possibilities. In addition to interactive documents, there are conversations to join and items to purchase. Not only is there a vast number of possibilities, but they vary widely in quality (Terveen and Hill, 2001).

The "information overload" and the fact that it is impossible to display all data on the server on the website and the desperate need of e-commerce sites, like that of Amazon.com, eBay.com and the like, to reach their customers in a more personalized (to reach their customers in order to increase their sales) way has forced for a system to be developed that can recommend items to make the life of user on the internet easy (Schafer et al, 2001).

Information overload

Information overload is a description given to the phenomenon where so much information is taken in by the human brain that it becomes nearly impossible to process it. The reason that information overload has become so prevalent in today's world is explained by the complexities

of the communications systems available to human beings. Instant communication is available by e-mail, cell phones, text messaging, and instant messaging. Added to that, are the thousands of academic journals on the Internet, and even more information freely distributed through blogs and amateur website and when all that is considered, information is being exchanged at rates never before experienced. Further, when considering these sources, much of the information being presented may not be original. As a result, many times the reader or viewer will spend time going over information already received from other sources. If there are discrepancies, or even small differences in the way the information is presented, this could create confusion, leading an individual to feel overwhelmed. This is a common symptom blamed on information overload (Toffler, 1981)

The “information overload” problem affects our everyday experience while searching for knowledge on a topic. To overcome this problem, we often rely on suggestions from others who have more experience on the topic. However, in Web case where there are numerous suggestions, it is not easy to detect the trustworthy ones. Shifting from individual to collective suggestions, the process of recommendation becomes controllable (Symeonidis et al, 2007).

To alleviate the information overload problem (especially on the Web), traditional Information retrieval techniques have been employed to assist the users in finding their information. One of the most widely applied information retrieval techniques for assisting the users in finding their information is the keyword based search, as adopted by many Web search engines. However, without prior knowledge of the retrieval process, or the keywords which accurately depict the search topic, discovering the desired information can be a tedious and formidable task (Haruechaiyasak et al, 2004).

Even though we are often making choices without sufficient personal experience of the alternatives that are available to us in different circumstances, in our everyday life, we sometimes rely on recommendations from other people either by word of mouth, recommendation letters or on movie and book reviews to select from the huge amount of Information that is available in different places but this suggestion is not enough in this digital age (Symeonidis et al, 2007).

2.1. Definition of Recommender System

A recommender system is an intermediary program (or an agent) with a user interface that automatically and intelligently generates a list of information which suits an individual's needs (Haruechaiyasak et al, 2004). It provides advice to users about items they might wish to purchase or examine. Recommendations made by such systems can help users navigate through large information spaces of product descriptions, news articles or other items (Burke, 2006).

When people have to make a choice without any personal knowledge of the alternatives from huge amount of options, a natural course of action is to rely on the experience and opinions of others. We seek recommendations from people who are familiar with the choices we face, who have been helpful in the past, whose perspectives we value, or who are recognized experts. We might turn to friends or colleagues, the owner of a neighborhood bookstore, movie reviews in a newspaper or magazine, or Consumers Union product ratings. And we may find the social process of meeting and conversing with people who share our interests as important as the recommendations we receive (Terveen and Hill, 2001).

Recommender systems intend to provide people with recommendations of products they will appreciate, based on their past preferences, history of purchase, and demographic information (Ziegler et al, 2005).

Recommender systems can now be found in many modern applications that expose the user to huge collections of items both commercially and in the research community, where many approaches have been suggested for providing recommendations. Such systems typically provide the user with a list of recommended items they might prefer, or predict how much they might prefer each item. It helps to match the users with items to make ease information overload and sales assistance (Shani and Asela, 2010). It is the web based application that uses the opinions of the community of users to help customers in the decision making and product selection process to identify content of interest from vast set of choices (Zanker, 2010). These choices may be books, music CDs, movie DVDs, flight tickets, hotel reservations, websites, field

of studies, and they may be people (e.g. find an expert, someone to follow) (Ning , 2007). In the figure below we can see how user’s web activity is full of important information that can be used by a recommender system for suggestion purpose.

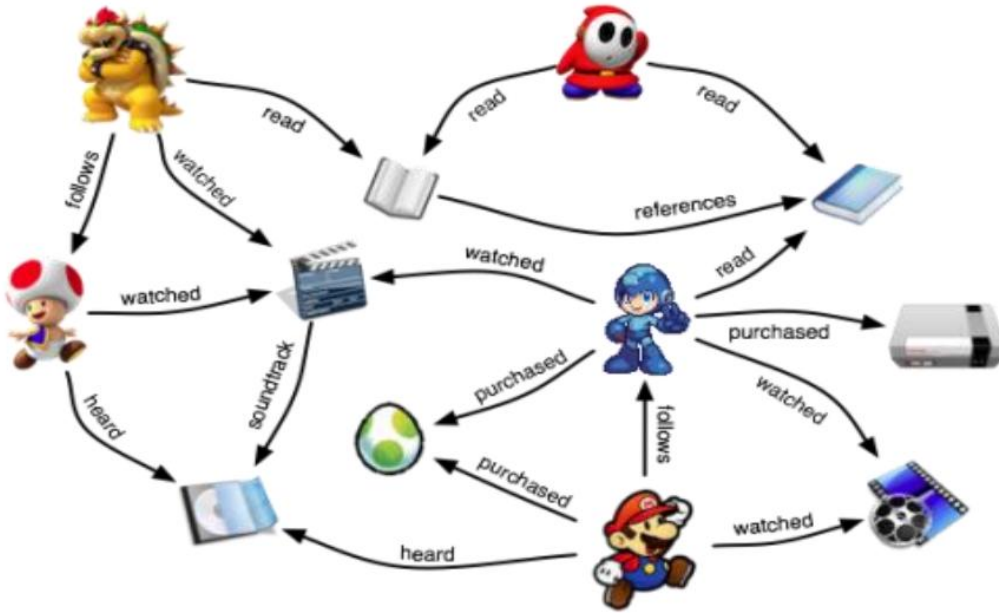


Figure 2. 1 User interactions on the web (source: Ning , 2007).

In a recommender system, the items of interest and the user preferences are represented in various forms, which may involve one or more variables. Particularly, in systems where recommendations are based on the opinion of others, it is crucial to incorporate the multiple criteria that affect the users’ opinions into the recommendation problem. Several recommender systems have already been engaging multiple criteria for the production of recommendations. Such systems, referred to as multi-criteria recommenders, early demonstrated the potential of applying multi-criteria decision making methods to facilitate recommendation in numerous application domains (Manouselis and Costopoulou , 2006).

2.2. Architecture for Recommender System

A recommendation seeker may ask for a recommendation, or a recommender may produce recommendations to the user with no prompting. Seekers may volunteer their own preferences, or recommenders may ask about them. Based on a set of known preferences

his/her own, the seeker's, and those of other people, often people who received recommendations in the past, the recommender recommends items the seeker probably will like. The seeker may use the recommendation to select items from the universe or to communicate with others (Terveen and Hill, 2001). The architecture of recommender system is represented in figure 2.2 below.

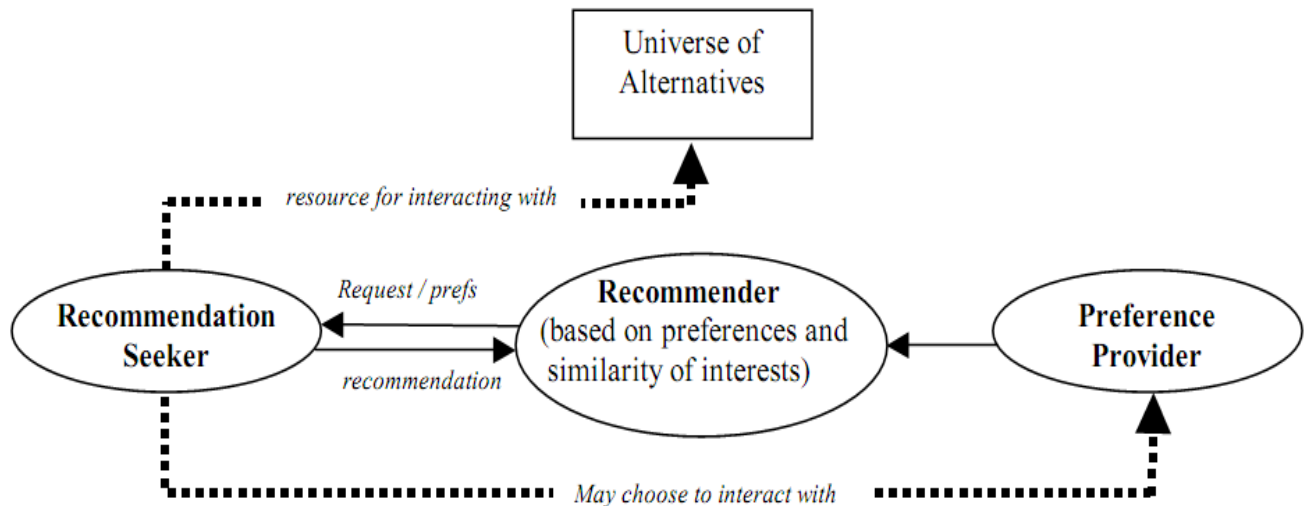


Figure 2. 2 Architecture of recommender system (source: Terveen and Hill, 2001)

As shown in the above figure, the recommendation seeker (students) asks recommendation with the help of recommender system from preference provider and universe of alternatives by entering the required queries by them. Based on the items listed in the preference provider, the system provides solution or recommendation by measuring similarity between new cases which is provided by the student and existing case stored in the case base. Finally, the system will provide best similar cases to the student from the existing cases.

In recommender systems, the utility of an item is usually represented by a rating, which measures how much a specific user is interested in the item. Depending on the application, the ratings can either be specified by the users, or be computed by the system. Each element of the user space U can be described with a profile that may include several demographic characteristics, such as gender, age, nationality, marital status, etc., and/or some information about the user's tastes, interests and preferences (Adomavicius & Tuzhilin, 2005).

The way in which such user profiles and item descriptions are defined is a key point in any recommender system. However, it is not the only factor that influences the efficiency and effectiveness of the recommendation processes (Adomavicius & Tuzhilin, 2005). Analogously, each element of the item space I may be described with a set of characteristics or features. For example, in a movie recommender system, movies can be described not only by their titles, but also by their genres, principal actors, etc.

As shown bellow, the basic components of a recommender system make interaction to give recommendation for the user. First, a user profile learning module (explicitly or implicitly) captures the preferences from the user. Once the system “knows” about the user’s tastes and interests, it compares and/or combines user profiles and item descriptions. Then item with a maximum gain to the user is recommended (Schafer et al, 2001).

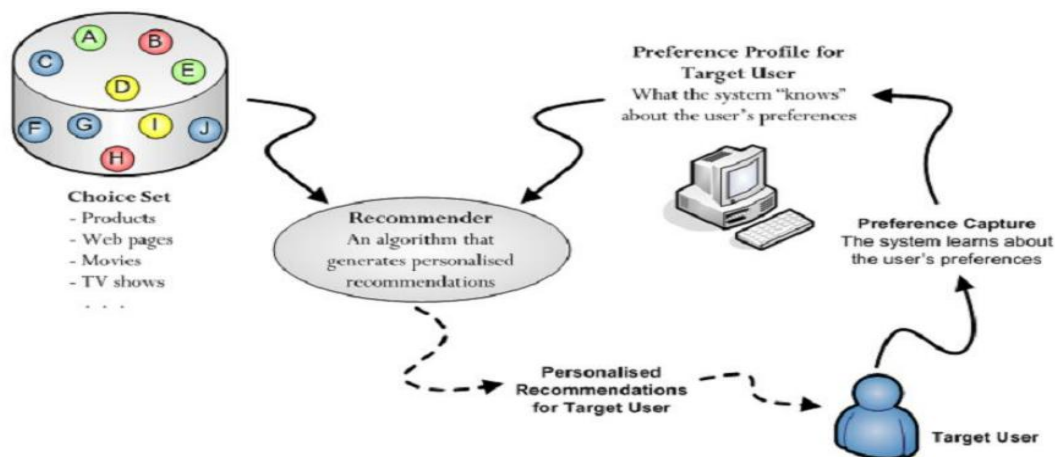


Figure 2. 3 Recommender System Process (source: Schafer et al, 2001)

2.3. Techniques applied in Recommender System

Recommender systems can be classified into the following categories, based on how recommendations are made (Adomavicius, 2005):

- **Content-based recommender systems**, in which the user is recommended items similar to those the user preferred in the past,

- **Knowledge based recommender system** in which it depends either on explicit domain knowledge about the items or knowledge about the users to derive relevant recommendation,
- **Collaborative filtering systems**, in which the user is recommended items that people with similar tastes and preferences liked in the past.

Due to the shortcomings proper of each of these strategies alone, combinations of both have been investigated in the so called **hybrid recommender systems**, empirically demonstrating their better effectiveness (Schafer et al., 2001).

The detail of each system will discussed as follows.

2.3.1. Content Based Recommender System

Content based recommender system recommend an item to a user based upon a description of the item and a profile of the user's interests (Pazzani and Billsus, 2007).

Content-based recommendation systems may be used in a variety of domains ranging from recommending web pages, news articles, restaurants, television programs, and items for sale. Although the details of various systems differ, content-based recommendation systems share in common a means for describing the items that may be recommended, a means for creating a profile of the user that describes the types of items the user likes, and a means of comparing items to the user profile to determine what to recommend. The profile is often created and updated automatically in response to feedback on the desirability of items that have been presented to the user (Pazzani and Billsus, 2007). The user will be recommended items similar to the ones user preferred in the past (Adomavicius, 2005). Content based recommender system recommends based on "show me more of the same as what I have liked" (Mobasher,2002).

Content based recommender system implement strategies for representing items and creates user profile that describes the types of items the user likes/dislikes. Each user is assumed to operate independently and items are represented by some features like Movies represented by actors, director, plot etc. The profile is often created and updated automatically in response to

feedback on the desirability of items that have been presented to the user. After that Filtering based on the comparison between the content (features) of the items and the user preferences as defined in the user profile can be made (Pazzani and Billsus, 2007).

In content-based recommendation methods, the utility $U(C, S)$ of item S for user C is estimated based on the utilities $U(C, S)$ assigned by user C to items $S_1 \dots S_n$ that are “similar” to item S . For example, in a movie recommendation application, in order to recommend movies to user C , the content-based recommender system tries to understand the commonalities among the movies user c has rated highly in the past (specific actors, directors, genres, subject matter, etc.). Then, only the movies that have a high degree of similarity to whatever the user’s preferences would be recommended (Semeraro, 2010).

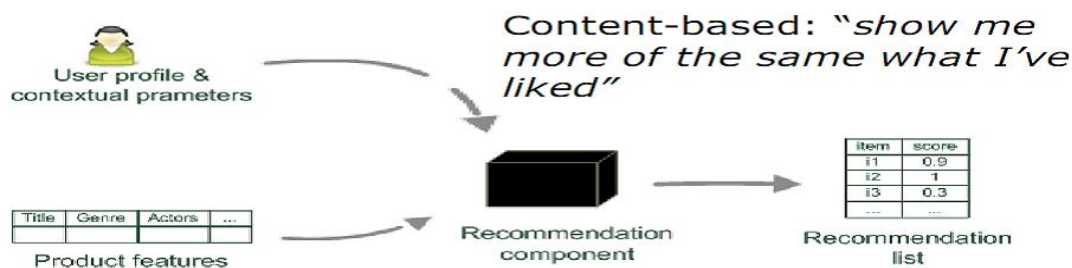


Figure 2. 4 Content based recommender system (source: Semeraro, 2010)

The content based approach to recommendation has its roots in information retrieval (Yates and Ribeiro, 1999) and information filtering (Belkin and Croft, 1992) research. Because of the significant and early advancements made by the information retrieval and filtering communities and because of the importance of several text-based applications, many current content-based systems focus on recommending items containing textual information, such as documents, Web sites (URLs), and Usenet news messages. The improvement over the traditional information retrieval approaches comes from the use of user profiles that contain information about users’ tastes, preferences, and needs. The profiling information can be elicited from users explicitly, e.g., through questionnaires, or implicitly learned from their transactional behavior over time (Cantador et al, 2010).

Item representation:

Items that can be recommended to the user are often stored in a database table that contains records (rows) and properties (columns) called attributes, characteristics, fields or variables in different publications. Each item in a database is described by the same set of attributes, and there is a known set of values that the attributes may have. In this case, many machine learning algorithms may be used to learn a user profile, or a menu interface can easily be created to allow a user to create a profile. The database may contain structured, semi-structured or unstructured data (Pazzani and Daniel Billsus, 2007).

User Profiles:

Profile of the user's interest is used by most recommendation systems in content based recommender system. This profile consists of a number of information (Pazzani and Daniel Billsus,2007).

- A. A model of the user's preferences. This is the description of the items that interest the user. There are many possible alternative representations of this description, but one common representation is a function that for any item predicts the likelihood that the user is interested in that item.
- B. A history of the user's interactions with the recommendation system. This may include storing the items that a user has viewed together with other information about the user's interaction.

2.3.1.1. Advantages of content based recommender system

According to Pazzani and Billsus (2007), content based recommender system has the following advantages.

User independent: Content based recommender system exploits solely ratings provided by the active user to build her own profile. No need for data on other users.

Transparency: content based recommender system can provide explanations for recommended items by listing content-features that caused an item to be recommended.

New item: Content based recommender system is capable of recommending new and unknown items.

2.3.1.2. Limitation of content based recommender system

Limited Content Analyses: Content-based recommendations are restricted by the features that are explicitly associated with the items to be recommended. For example movie recommendation system may be only based on written materials about the author like name, title, etc (Gutiérrez, 2008).

Over specialization: When the system can only recommend items that score highly against a user's profile, the systems recommendation will be predictable but most users like recommenders that recommend unexpected items (Shoham and Balabanovic, 1997).

New user problem: new user, having very few ratings, would not be able to get accurate recommendations (Adomavicius, 2005).

2.3.2. Collaborative Filtering Recommender System

Collaborative filtering has been defined by Chau et al , (2002) as “a collaboration in which people help one another perform filtering by recording their reactions to documents they read.” In collaborative filtering, a user's actions and analyses regarding a particular piece of information are recorded for the benefit of a larger community; i.e., potential users of the same information. Members of the community can get benefit from others' experience before deciding to use new information. In essence, collaborative filtering systems essentially automate the process of “word-of-mouth” recommendations (Shardanand and Maes, 1995).

The idea of collaborative filtering is in finding users in a community that share appreciations (York et al, 2003). If two users have same or almost same rated items in common, then they have similar tastes. Such users build a group or a so called neighborhood. A user gets recommendations to those items that he/she hasn't rated before, but that were already positively rated by users in his/her neighborhood. For example, the following figure shows that all three users rate the movies positively and with similar marks. That means they have similar

taste and build a neighborhood. The user A hasn't rated the movie "TRON: Legacy", which probably mean that he hasn't watched it yet. As the movie was positively rated by the other users, A will get this item recommended.

Movies Users	Titanic	Gladiator	Black Swan	The Fighter	TRON: Legacy
A	8	7	9	10	-
B	9	7	9	9	10
C	9	8	9	8	9

Figure 2. 5 Collaborative Filtering Recommender System. (Source: Asanov, 2011).

Collaborative recommender systems care about the taste of user. The taste is considered to be constant or at least change slowly. We can distinguish two popular approaches of collaborative filtering recommender system (Sun and Luo, 2010). **User-based approach** which is the users performs the main role. If certain majority of the customers has the same taste, then they join into one group. Recommendations are given to user based on evaluation of items by other users from the same group with whom he/she shares common preferences (Sun and Luo, 2010). **Item-based approach** Referring to the fact that the taste of users remains constant or change very slightly similar items build neighborhoods based on appreciations of users. Afterwards the system generates recommendations with items in the neighborhood that a user would prefer (Sun and Luo, 2010).

Unlike content-based recommendation methods, collaborative recommender systems (or collaborative filtering systems) try to predict the utility of items for a particular user based on the items previously rated by other users (Semeraro, 2010). It recommends based on the principle "Tell me what is popular among my peers" (Gutiérrez, 2008).

As indicated by Herlocker (2000), collaborative filtering systems work by collecting human judgments (known as rating) for items in a given domain and matching together people who share the same information needs or the same tastes.

The goal of collaborative filtering is to predict the preferences of one user, referred to as the active user based on the preferences of a group of users of similar interest (Goldberg et al, 2000). For example, given the active users ratings for several documents and a database of other users' ratings, the system predicts how the active user would rate unseen documents. The key idea is that the active user will prefer those items that likeminded people prefer. The effectiveness of any collaborative filtering systems is ultimately predicted on the underlying assumption that human preferences are correlated (Pennock and Horvitz, 2000).

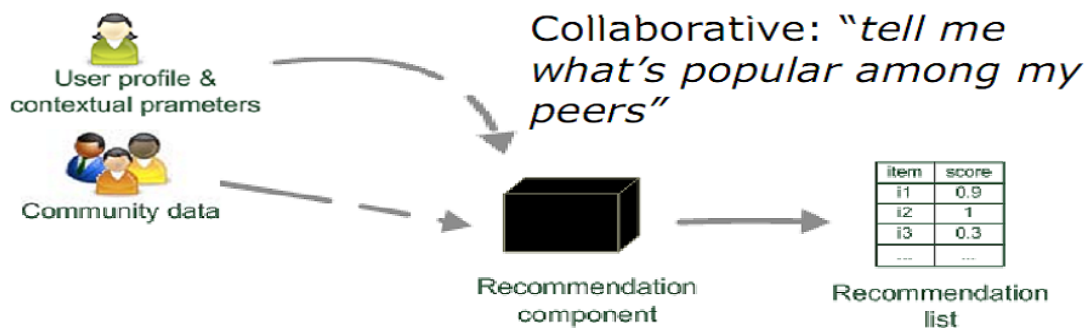


Figure 2. 6 Collaborative Filtering Recommender System (source: Mobasher,2002)

The utility $u(c; s)$ of item s for user c is estimated based on the utilities $u(c_j; s)$ assigned to item s by those users $c_j \in C$ who are "similar" to user c . for example, in movie recommendation application, in order to recommend movies to user c , the collaborative recommender system tries to find the "peers" of user c , i.e., other users that have similar tastes in movies (rate the same movies similarly). Then, only the movies that are most liked by the "peers" of user c would be recommended.

2.3.2.1. Uses for Collaborative Filtering

Early collaborative filtering systems were designed to explicitly provide users with information about items. That is, users visited a website for the purpose of receiving recommendations from the collaborative filtering system. Later, websites began to use collaborative filtering systems behind the scenes to adapt their content to users, such as choosing which news articles a website should be presenting prominently to the user (Schafer et al., 2007).

Tasks for which people use collaborative filtering that have been studied include (Schafer et al, 2007).

- **Help me find new items I might like.** In a world of information overload, evaluating all things is difficult. This system presents a few for the user to choose from. This has been applied most commonly to consumer items (music, books, movies), but may also be applied to research papers, web pages, or other ratable items.
- **Advise me on a particular item.** We have a particular item in mind; does the community know whether it is good or bad?
- **Help me find a user (or some users) I might like.** Sometimes, knowing who to focus on is as important as knowing what to focus on.
- **Help our group find something new that we might like.** Collaborative filtering can help groups of people find items that maximize value to group as a whole (Cosley et al, 2001). For example, a couple that wishes to see a movie together or a research group that wishes to read an appropriate paper.
- **Help me find a mixture of “new” and “old” items.** I might wish a “balanced diet” of restaurants, including ones I have eaten in previously; or, I might wish to go to a restaurant with a group of people, even if some have already been there; or, I might wish to purchase some groceries that are appropriate for my shopping cart, even if I have already bought them before.

2.3.2.2. Advantage of collaborative filtering recommender system

The main advantages of this type of recommender systems is that they are more effective when it comes to customer satisfaction as they recommend the most appropriate items to users which are personalized at the same time. The collaborative filtering algorithms are designed such that the accuracy of their prediction increases tremendously over time as more user preferences are added to the database, irrespective of the size of the database (Aradhya, 2001).

2.3.2.3. Disadvantage of collaborative filtering recommender system

- **New User Problem:** It has the same problem as with content-based system. In order to make accurate recommendations, the system must first learn the user's preferences from the ratings that the user gives.
- **New Item Problem:** New items are added regularly to recommender systems. Collaborative systems rely solely on users' preferences to make recommendations. Therefore, until the new item is rated by a substantial number of users, the recommender system would not be able to recommend it.

2.3.3. Knowledge based recommender systems

Knowledge includes facts about the real world entities and the relationship between them. It is an understanding gained through experience and familiarity with the way to perform a task. It is an accumulation of facts, procedural rules, or heuristics (Russel and Norvig, 2010).

Knowledge base is used to store facts and rules about real world. In order to solve problems, the computer needs an internal model of the world. This model contains facts and rules, for example, the description of relevant objects and the relations between these objects. All information must be stored in such a way that it is readily accessible called knowledge base (Russel and Norvig, 2010). The system will extract this stored knowledge from knowledge base to give recommendation to the user.

A knowledge based recommender reasons about the fit between a user's need and the features of available products (Burke, 2000).

Knowledge based recommender system uses knowledge about users and products to pursue knowledge based approach to generate a recommendation, reasoning about what products meet the user's requirements. The Personal Logic recommender system offers a dialog that effectively walks the user down a discrimination tree of product features (Burke, 2002). It depends either on explicit domain knowledge about the items or knowledge about the users to

derive relevant recommendation. This system works based on the principle “tell me what fits based on my needs” (Mobasher, 2002).

A knowledge based recommender system avoids some of the drawbacks of collaborative filtering and content based recommender system. It does not have a ramp-up problem since its recommendations do not depend on a base of user ratings. Gathering information about a particular user is not required because its judgments are independent of individual tastes. This makes knowledge based recommenders not only valuable systems on their own, but also highly complementary to other recommender systems (Burke, 2002).

2.3.3.1. Advantage of knowledge base recommender system

As stated by Jannach et al. (2011), knowledge based recommender system have its own advantages and disadvantages. Some of advantages are:

- It is deterministic recommendations
- Assured quality
- There is no cold start
- It can resemble sales dialogue

2.3.3.2. Disadvantage of knowledge base recommender system

Some of advantages of knowledge based recommender system are:

- Knowledge engineering effort to Bootstrap
- Suggestion ability is basically static
- It does not react to short-term trends

2.3.4. Hybrid Recommender System

Disadvantages of the Collaborative Filtering and Content Based approaches can be solved by combining the two into a hybrid method. Many hybrids approaches use two recommendation algorithms and combine their results in some manner, such as combining the results by their relevance, mixing the output of the two algorithms, switching from content based into

collaborative once the cold start phase is over, or using the output of one algorithm as input to the second algorithm. Using hybrid approaches we can avoid some limitations and problems of other recommender systems, like the cold-start problem (Burke, 2002). The combination of approaches can proceed in different ways (Adomavicius, 2005).

- Separate implementation of algorithms and joining the results.
- Utilize some rules of content-based filtering in collaborative approach.
- Utilize some rules of collaborative filtering in content based approach.
- Create a unified recommender system that brings together both approaches.

Some of the hybrid recommenders according to Buke (2002) are:

Weighted recommenders:

The results of all the recommendation techniques are used to compute the score of an item which is to be recommended. The assumption used in this technique is that relative value of numerous other techniques is uniform across the space of possible items.

Switching recommenders:

This hybrid incorporates the item level sensitivity to its strategy using which, the recommender switches between the techniques used for recommendation.

The notable characteristic of this system is that it is sensitive to the strength and weakness of the constituent recommenders.

Mixed recommenders:

These types of recommenders are used when there is a need to make large number of recommendations in a short period of time. This type of recommender takes suggestions from multiple sources to generate final results.

2.4. Process of Recommender System

Most recommender systems work in a sequential manner. The sequential nature of the recommendation process, where at each stage a new list is calculated based on the past ratings, is referred to as a sequential optimization process. The other sequential aspect of the

recommendation process is called an optimal recommendation which depends not only on previous items selected, but also on the order in which those items are selected Guy et al., (2005) and Zimdars et al., (2001) recognized this possible dependency and suggested the use of an auto-regressive model to represent it. They divided a sequence of transactions X_1, \dots, X_T (for example, field of study selection, web-page views, etc) into cases $(X_{t-k}, \dots, X_{t-1}, X_t)$ for $t = 1, \dots, T$. They then built a model to predict the last column given the other columns, under the assumption that the cases were exchangeable.

According to Vozalis and Margaritis, (2001), the two basic entities which appear in any Recommender System are the user / customer and the item / product. A user is a person who utilizes the Recommender System providing his opinion about various items and receives recommendations about new items from the system.

2.5. Input of Recommender System

The input to a Recommender System depends on the type of the employed filtering algorithm. Inputs belong to one of the following categories (Vozalis and Margaritis, 2001).

Ratings: which express the opinion of users on items. Ratings are normally provided by the user and follow a specified numerical scale. A common rating scheme is the binary rating scheme, which allows only ratings of either 0 or 1. Ratings can also be gathered implicitly from the user's purchase history, web logs, hyperlink visits, browsing habits or other types of information access patterns.

Demographic data: Which refer to information such as the age, the gender and the education of the users. This kind of data is usually difficult to obtain. It is normally collected explicitly from the user.

Content data: Which are based on a textual analysis of documents related to the items rated by the user. The features extracted by this analysis are used as input to the Filtering algorithm in order to infer a user profile.

Cases: Case based recommendation requires that content is represented in a more structured way. Each item is stored as a case. A case could be thought of as a record in a database which is a collection of features or attributes and their values.

The goal of Recommender Systems is to generate suggestions about new items or to predict the utility of a specific item for a particular user. In both cases the process is based on the input provided, which is related to the preferences of that user.

2.6. Output of a Recommender System

The output of a Recommender System can be either a Prediction or a Recommendation (Vozalis and Margaritis, 2001).

A **Prediction** is expressed as a numerical value, $r_{a,j}$, which represents the anticipated opinion of active user u_a for item i_j . This predicted value should necessarily be within the same numerical scale as the input referring to the opinions provided initially by active user u_a . This form of Recommender Systems output is also known as Individual Scoring.

A **Recommendation** is expressed as a list of N items, where $N \leq n$, (where n items in knowledge base) which the active user is expected to like the most. The usual approach in that case requires this list to include only items that the active user has not already purchased viewed or rated. This form of Recommendation output is also known as Top-N Recommendation or Ranked Scoring.

2.7. Case Based Recommendation

Case based recommendation is the form of content based filtering and relies on product data being stored in a structured manner (Barry 2009).

Case Based Recommendation is one of the most successful machine learning methodologies that exploit a knowledge rich representation of the application domain. It is a problem solving methodology that addresses a new problem by first retrieving a past, already solved similar case, and then reusing that case for solving the current problem. In the most simple application

of case base recommender system to recommendation generation, the case base models, the products to be recommended and the set of suggested/recommended products are retrieved from the case base by searching for products similar to that partially described by the user (Fabiana and Francesco, 2003).

2.8. Knowledge Representation

Since knowledge is the description of the world, it needs knowledge representation in the way that knowledge is encoded. Different types of knowledge require different kinds of representation. Knowledge Representation expresses knowledge explicitly in a computer tractable way such that the agent can reason out (Chakraborty, 2010).

Knowledge representation is most fundamentally, a representative, a substitute for the thing itself, enables an entity to determine consequences by reasoning about the world rather than taking action on it. It is a medium of human expression, a language in which we say things about the world. It is a set of ontological commitment i.e, an answer to the question; in what terms should I think about the world (Randall et al, 1993).

To build knowledge base, several representation methods/models can be used. The most widely used knowledge representation techniques includes semantic network, logics, rules, case base and frames (Chakraborty, 2010).

Semantic networks: - are directed graph used for describing semantic relationships between knowledge items in a knowledge base. Semantic network involves nodes and links between nodes. The node represents objects and the link represents relations between nodes. The network defines a set of binary relations on a set of nodes (Meskerem, 2009).

Logics:- Logic is a formal and declarative language in which knowledge can be represented such that conclusions can easily be drawn. It contains syntaxes and semantics. Syntax refers what expressions/structures are allowed in the language.

Example:

mycar (red) is ok, but mycar(grey or green) is not. Semantics expresses what sentences mean, in terms of a mapping to real world. Semantics relate sentences to reality. From the previous example mycar (red) means that my car is red.

Frames: - A frame is collection of slots (attributes) that characterizes an object. Each slot may be filled with another frame, procedure or a value. A slot includes details of each data object provide links to other frame, containing procedural code, linking to other applications to obtain data or written data.

Rules and Facts

Rules: - A rule based system consists of if-then rules, a bunch of facts, and an interpreter controlling the application of the rules, given the facts.

These if-then rule statements are used to formulate the conditional statements that comprise the complete knowledge base. A single if-then rule assumes the form

“if x is A then y is B” and the if part of the rule ‘x is A’ is called the antecedent or premise, while the then part of the rule ‘y is B’ is called the consequent or conclusion (Abraham, 2005).

The IF portion of a rule is a condition, also called a premise or an antecedent, which tests the truth value of set of facts. If these are found true, THEN portion of a rule also called the action, the conclusion or the consequent is inferred as a new set of facts (Chakraborty, 2010).

As stated by Abraham, there are two broad kinds of inference engines used in rule based systems to create knowledge base system.

Forward chaining: In a forward chaining system, the initial facts are processed first, and keep using the rules to draw new conclusions given those facts. Forward chaining systems are primarily data driven.

Backward chaining: In a backward chaining system, the hypothesis or goal we are trying to reach is processed first, and keep looking for rules that would allows concluding that hypothesis. Backward chaining systems are goal-driven.

For example,

Rule 1: If A and C then Y

Rule 2: If A and X then Z

Rule 3: If B then X

Rule 4: If Z then D

If the task is to prove that D is true, given A and B are true. According to forward chaining, start with Rule 1 and go on downward. Rule 3 is the only one that fires in the first iteration. After the first iteration, it can be concluded that A, B, and X are true. The second iteration uses this valuable information. After the second iteration, Rule 2 fires adding Z is true, which in turn helps Rule 4 to fire, proving that D is true.

In the backward chaining method, processing starts with the desired goal, and then attempts to find evidence for proving the goal. From the above example, the task to prove that D is true would be initiated by first finding a rule that proves D. Rule 4 does so, which also provides a sub-goal to prove that Z is true. Then Rule 2 comes into play, and as it is already known that A is true, the new sub-goal is to show that X is true. Rule 3 provides the next sub-goal of proving that B is true. But that B is true is one of the given assertions. Therefore, it could be concluded that X is true, which implies that Z is true, which in turn also implies that D is true in turn also implies that D is true.

Facts: - facts are data or instances that are specific and unique. The known information/data about the problem is represented in facts or objects. These facts or objects are used to trigger left hand side conditions of rules. The set of current facts or objects is dynamic; it changes throughout the execution of the program as new information is provided by users or as a result of rules that have been fired.

Case base: - Case-based means using old experiences to understand and solve new problems. In case based reasoning, a reasoner remembers a previous situation similar to the current one and uses that to solve the new problem. Case based reasoning can mean adapting old solutions to meet new demands; using old cases to explain new situations; using old cases to critique new solutions; or reasoning from precedents to interpret a new situation

(much like lawyers do) or create an equitable solution to a new problem (L. Kolodne, 1992).

For this study knowledge (cases) used for decision making at the time of field of study selection process has been collected from students, relevant documents from ministry of education (national agency for examination). To collect relevant attributes and case, semi-structured and structured interview were used. Such cases are very important inputs to recommender system for the appropriate decision of field of study selection process.

2.9. Student placement process

Information about courses, curriculum, research and facilities in the field of education is mainly available on the web. Information about educational institution with richer description enables the students, parents and people to make their choice more efficiently and scientifically (Satyanarayana and Rajagoplan, 2007).

A university education has become a basic part of most people's preparation for working life. University admission is a complex decision process that goes beyond simply matching test scores and admission requirements. Admission to university is therefore a topic of importance. How a student chooses a university, and conversely how a university chooses a student, determines the success of both sides in carrying through the education (Fong and Biuk-Aghai, 2009). One of the main problems faced by university students is to take the right decision in relation to their academic program based on available information (for example field of study, course, section, class room and etc) (Vialardi et al, 2009).

2.9.1. Student`s Placement Process in Ethiopia

When we come to Ethiopia, as described in the introduction part, a student join higher education when they completes their preparatory education (Grade 11-12) and gets at least the minimum entrance examination score.

Those Students who fulfill the requirement of higher education entrance score can join both governmental and private higher education institution. However, to join governmental

higher education institution, students should pass through the placement process of the Ministry of Education. The Ministry of Education (National Agency for Examination) has placement policy to manage this placement process.

Based on the criteria set by the Ministry of Education, those students who satisfy the entrance requirement will fill in a form about their choice of field of study and universities they would like to join.

As stated by some preparatory school directors, students are required to choose their field of study and university in order to succeed in their higher education. They are provided with the necessary orientation manual to do this; - but some of the students do not select their field of study and university with confidence. After completing their form, these students get upset with their wrong choose and ask their schools to allow them to change their choices. The schools are forced to do so again and again to help students to succeed with their higher education.

The placement process will be done as follows (Getachew Feleke, 2008).

- List fields of studies at each universities and colleges throughout the country;
- Set the maximum number of students a university or a college can accept under each specific field of study per year;
- Encode students score and choice information in one file which is collected from preparatory schools;
- Select one field of study from the list given that the total number of students who can be placed in that field is known and Sort those who have chosen the specific field of study as their first choice in descending order based on their score;
- Evaluate the cases of each student with disabilities and special cases, and if their case is acceptable assign those students with their choice of field of study and as well as university.
- Sort only female students (government policy of affirmative action) who have chosen that specific field of study as their first choice in descending order based on their score so that female students get priority.

- If the number of female students who have chosen that field as their first choice is greater than one fourth ($1/4$) of the available capacity of institutions only the top students who are within the range of quarter of total capacity will get that specific field of study and the remaining female students compete with all students;
- If the number of students who have chosen that field is greater than the available capacity of universities, only the top students who are within the range will get that specific field of study, the rest will compete with their next choice and so on.
- List students who are not assigned to a certain field of study and whose second choice is that specific field of study in descending order of their score and assign those students within the range of current field of study;
- Repeat these steps until all students are assigned to specific field of study;
- If there are some students who are not assigned to a certain field of study, assign those students in the available fields of study based on their entrance exam result.

These processes are to assign students who made correct choice of field of study and university.

Finally, all students are assigned to a certain field of study at a certain higher education institution, so the next activity is announcing the placement information to students, schools, and higher institutions. Even though, it has constant problems especially to the rural area students who have no chance to get internet access, the Ministry is doing this placement announcement using National Agency of Examination website. Students can see their field of study and university placement result by login with their name identification number and school plasma TV.

2.10. Related works

There seems to be precious little directly related works to be found. This is because the educational process differs at different countries. For instance, in Ethiopian higher education institutions, students' placement can be processed with Ministry of Education. Some of the related works can be discussed below to assess what others follow and which one is used as a model for our country.

According to Ekdhal et al, (2002), students are required to choose the courses they are interested in reading for the subsequent semesters. It is the student's responsibility to make sure that she/he gets the courses needed to graduate. If the choice is not made, students will not be allowed to take that course. It is also student's responsibility to make sure that she/he has read for their prerequisite courses for their choice, as well as to make sure that the selected courses are all to simultaneously be part of a degree (Ekdhal et al, 2002). It is quite that students have a lot to take into consideration before choosing courses, and that it can be hard to get a good overview of all available courses. The student course recommender is meant to aid students in this task. As stated by Ekdhal, the initial goal of the system was to utilize the machine learning system to estimate user interests in courses given at the university. The result of the research was the system that was implemented using standard client/ server architecture. The client, a simple graphical interface to the server, contains the logic and parts needed to handle client request for course recommendation (Ekdhal et al, 2002). According to the summary of Ekdhal, students need recommender system to advice them in the course selection process. But in Ethiopia, rather than course recommendation field of study recommendation is very important since the success of students highly dependent of their interest of field of study.

The other related work is "An Automated University Admission Recommender System for Secondary School Students" by Fong and Aghai, (2009).

Admission and placement of students is based on the perspective of Universities who knows little about the incoming student background but not based on the perspective of high schools who knows the detail of their students. There is value in extending the university admission process to include secondary schools (Fong and Aghai, 2009).

In this work, the author proposes a novel design of a recommender system that can provide recommendations about which universities a student should apply to, taking not only the student's secondary school scores but also other factors such as background interest and special skill into account.

In the summary of the author, education systems which do not have a standardized open exam for university admissions face the challenges of matching the right secondary school students with the right universities and field of studies and the ways that they should enter. This implies some manual processes are needed and web based recommendation system is very important for decision making. To do that, the author applied a hybrid data mining model to implement a recommender system prototype and analyze different data from secondary schools.

In addition to that, a work entitled “Recommender system for higher education” has been done to discuss the process of developing recommender system for educational institutions. The system is web based application that guides students for decision making based on their personal test. Information about course, curriculum research and facilities in the field of education is important to be available on the web. Clear information about educational activities with description enables students, partners and people to choice more efficiently and scientifically to make right decision (Satyanarayana and Rajagoplan, 2007).

The author concludes that, we can make beneficial use of artificial intelligence techniques like database design and selection, content based recommendations, user profiling, integrating groups of users with similar interests and integrating the domain knowledge and expertise. Hybrid recommendation system approach is important in educational institutions (Satyanarayana and Rajagoplan, 2007).

Bendakir and imeur (2006), tried to discuss about course recommendation using data mining techniques called association rule. According to the author, students often need guidance in choosing adequate courses to complete their academic degrees. Course recommender systems have been suggested in the literature as a tool to help students make informed course selections.

Students who join higher education degrees are faced with two main challenges: a myriad of courses or field of studies from which to choose, and a lack of knowledge about which courses or field of studies are relevant to follow and in what sequence. Mostly, it is according to their friends and colleagues’ manual recommendations that the majorities of them choose their field

of study and register for it. It would be useful to help students in finding courses of interest by the intermediary of web based recommender system (Bendakir and imeur, 2006).

The main focus of the author is on the effectiveness of the incorporation of data mining in course recommendation. The system is based on the following collaborative filtering algorithms: user-based and item-based (discussed earlier). According to the author, the system can predict the usefulness of courses to a particular student based on other users' course ratings. To get accurate recommendations, one must evaluate as many courses as possible.

Based on the evaluation results, the author suggests C4.5 as the best algorithm for course recommendation. The system cannot predict recommendations for students who have not taken any courses at the University.

Generally, there are many recommender systems developed globally. The study focus on students related works such as course recommender system, higher education recommender system, automated placement recommender system and etc.

When we come to our country Ethiopia, as far as my knowledge, there is no field of study recommender system developed yet. The only related work is "Student Advice Expert system" by Ebrahim Kassa, (1999). The author tried to develop expert systems that can advice students in their field of study selection process. But the system focuses only on Addis Ababa University evening students. This system guides students to make right decision when they choose field of study at evening program for all schools of Addis Ababa University.

Generally, developing recommender system for field of study selection is important for both students as well as National Agency for Examination. This work will be used as an input for further development of field of study selection related works.

CHAPTER THREE

KNOWLEDGE ACQUISITION AND MODELING

Industries and societies are becoming knowledge oriented and dependent of decision making ability of expert. Knowledge base system can act as an expert on demand; can save money by leveraging experts; allowing users to function at higher level and promoting consistency (Sajja & Akerkar, 2010).

3.1. Overview Knowledge Base System

The concept of knowledge base system is derived from the field of artificial intelligence. It is one of the main family members of the artificial intelligence. Artificial intelligence intends understanding of human intelligence and building of computer programs that are capable of simulation or acting one or more of intelligent behavior. Intelligent behavior includes cognitive skill like thinking, problem solving, learning understanding, emotion, consciousness, intuition and creativity, etc. Nowadays, some of the behaviors such as problem solving, learning and understanding are handled by computer programs (Sajja and Akerka, 2010).

Knowledge base system is a computer program that can solve problems like human expert by using knowledge about the application domain and problem solving techniques (Speel et al, 2001). Human experts use their knowledge about the domain and techniques that lead how to use the knowledge to solve problems. Computer knowledge base systems handle problems in the same way as humans do. The system represents knowledge about a specific application domain and uses one or more techniques that guide to use knowledge to solve problems (Sajja and Akerka, 2010).

Knowledge base system is the general term used for the process of eliciting, structuring and representing knowledge from some knowledge source mostly from human experts and developing a computational problem solving model, specifically a program to be used in some consultative or advisory role (Sagheb-Tehrani, 2009, Fredlund et al, 1996). Such programs are generally called expert system.

Industries and societies are becoming knowledge oriented and dependent of decision making ability of expert. Knowledge base system can act as an expert on demand; can save money by leveraging experts; allowing users to function at higher level and promoting consistency (Sajja & Akerkar, 2010).

Knowledge base systems increase productivity, document when there is shortage of knowledge for future use, enhances problem solving capability and this leads to increase quality in problem solving process.

As discussed by Sajja & Akerkar, 2010, Knowledge based systems are more useful than the traditional computer based information systems when:-

- There is shortage of experts
- Expertise is to be multiplied and stored for future use.
- There is more group of platform than one experts' knowledge.
- Intelligent assistance important for decision making.

However, the scarcity and nature of knowledge make the knowledge base system development process difficult and complex. Some of the limitations of knowledge base system are due to the following reasons (Sajja & Akerkar, 2010).

- Large volume of Knowledge Acquisition, representation and manipulation.
- Limitations of cognitive science and other scientific methods.
- Abstract nature of the knowledge.

3.2. Architecture of a Knowledge Base System

The knowledge base system consists of a Knowledge Base and a search program called Inference Engine. Figure shows 3.1 the main building blocks of knowledge base system (Sajja and Akerka, 2010).

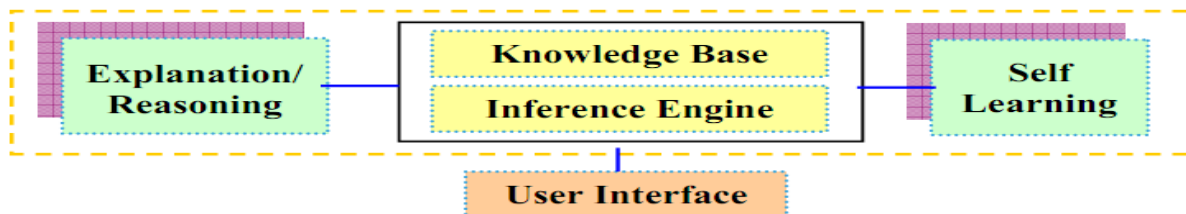


Figure 3. 1 Architecture of knowledge base system (source: Sajja and Akerka, 2010)

Inference engine: - also known as rule interpreter. It is a software program and problem solving component, which infers the knowledge available in the knowledge base. It allows new inferences to be made from the case specific data and the knowledge in the knowledge base.

Explanation/reasoning: - analyzes the structure of the reasoning performed by the system and explains it to the user.

Domain expert: - individuals who currently are experts in solving problem of the domain area.

User: - individuals who consults with the system to get advice with the system which would have been provided by the expert.

Knowledge base: - contains necessary knowledge about the domain that is required to handle problems. Knowledge can acquire from expert, documents, books, novice and others. There are many techniques to represent knowledge such as rules, semantic network, and frames.

3.3. Knowledge Base System Development

The knowledge base system development process is composed of the different components as illustrated in the following figure 3.2 (Sajja and Akerka, 2010).

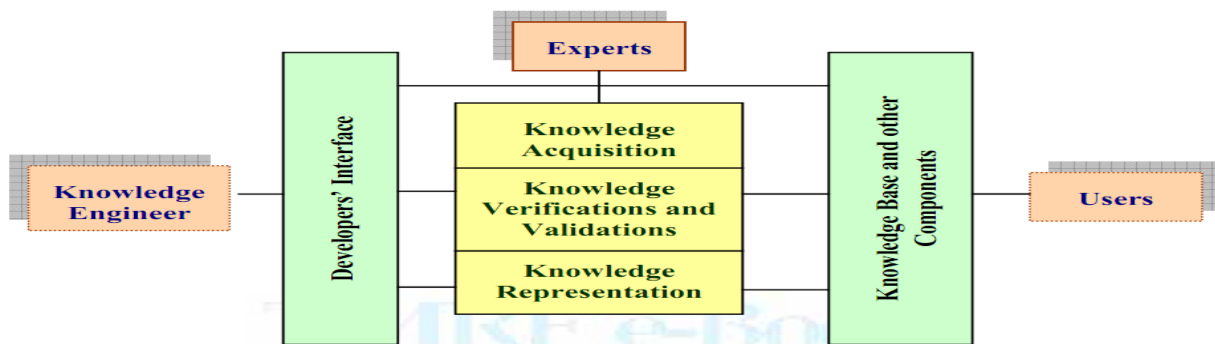


Figure 3. 2 Development of knowledge base system (source: Sajja and Akerka, 2010)

An expert is the one who has stored knowledge in his mind in a very abstract way. But every expert may not be familiar with knowledge based systems and the way how to develop an intelligent system.

Knowledge engineer is an individual who encode the tacit knowledge of an expert including related documents in declaration form that can be used by the knowledge base system (Meskerem, 2009).

3.3.1. Knowledge Acquisition

Knowledge acquisition is the most important process and vital stage in knowledge based system development. How knowledge is obtained and where it is obtained determines the usefulness of the system. Knowledge acquisition process has many steps. Some of them are: selecting a problem to be solved by the program, interviewing an expert, questionnaires, observation, record reviews, codifying the knowledge in some representation language, and refining the knowledge base by testing it and extending its capability (Clancey, 1984). This process is generally called knowledge elicitation.

Knowledge elicitation methods can be classified in many ways. Direct and indirect elicitation methods are the most common ways of knowledge elicitation. This shows how the knowledge engineers obtain information and knowledge from domain experts. Direct questioning the domain expert on how they are doing their job and how they can succeed is called direct elicitation. Where as in the indirect elicitation method the required information from domain expert is not requested directly but can be elicited by analyzing the session to extract the required information (Shadrick et al, 2005).

In this research, for the development of recommender system, there was different knowledge acquisition methods used such as:-

- Eliciting data and information from the domain expert
- Acquiring knowledge from relevant documents
- Acquiring knowledge from Ministry of Education
- Interpreting the data and information and making conclusion about the expert's knowledge that will be used as input for the recommender system
- Constructing model which describes the expert's knowledge.

3.3.1.1. Knowledge Acquisition from Relevant Document

In order to elicit knowledge from documents, relevant documents which are related to field of study selection process have been review. Document analysis collects information from existing documents. These documents includes promotional literature, brochures, manuals, employee hand books, reports, glossaries, course texts and training materials (Clark et al, 2008).

The documentary sources used in this study are:

- Articles that are published in different journals
- Vouchers
- Manuals, Guideline and forms that are used in the process of field of study selection
- Academic Websites related to field of study selection specially Ministry of Education and National Agency for Examination websites.

In addition to these, different variables that are considered in recommender system have been identified. There are documents in the preparatory schools which are not formal manuals that are used as guidance in the process of field of study selection. These documents have been reviewed to get knowledge for the system. For instance students background such as social science and natural science field of study have been collected from preparatory schools and bands and sub-bands which represents field and sub-fields of study were collected from the manual of National Agency for Examination.

3.3.1.2. Knowledge Acquisition from Domain Expert

Interview is one of the knowledge elicitation techniques which involve asking the domain expert on how they perform their task and become successful. Interview can be structured, semi structured and unstructured (Burge, 1998).

An interview with a set of predefined questions that can be answered by the expert is called structured interview which are well organized questions. Semi-structured interview is an interview with both open-ended and closed-ended questions that contains a guide. It is more

flexible than structured one because the interviewer has the chance to change the order of questions according to the context of response. Unstructured interview is a kind of interview that is not predetermined. It depends on the interaction of the researcher and domain expert (Clark et al, 2008).

Since domain expert is the individual who is expert in solving problem of domain area, the knowledge of experts play a vital role in structuring the knowledge base and developing the knowledge base system. When the knowledge base system elicited from expert increases the recommendation system for decision making will also increase. As a result of this, expert on the different profession such as head of departments and officers have been interviewed to obtain the needed knowledge for recommender system. The researcher interviewed all 32 (thirty two) department heads in Debre Markos University to collect relevant attributes. These department heads have their own believes to assign students to the respective bands and sub-bands. These departments can represent field of studies since students in the department are within that field of studies called bands and sub-bands. Therefore, the researcher used Semi-structured interview questions and collected relevant attributes from each field of studies with purposive data collecting technique. About thirteen attributes namely Preparatory class attend, special skill, interest1, interest2, interest3, background interest, family interest, preparatory best result, entrance exam result, health status, language dependency, duration and past experience were identified in such a way. After collecting and organizing main attributes from domain experts, the next step was collecting cases to represent case base. All the cases were collected from each department of Debre Markos University. The researcher used purposive data collecting techniques since there are targeted students who succeeded in their field of study selection. These successful students were selected from each department by the researcher to fill the necessary data within well structured interview questions. One hundred and five cases were collected from as many successful students and represented as case base that are used as previously solved cases.

3.3.1.3. Knowledge Acquisition from Ministry of Education

Domain experts from National Agency for Examination were also the main targeted groups to collect relevant attributes. Variables that are identified from documents and website of the agency that are used in field of study selection process such as lists of field of study, sub-lists under each field of study were the main variables.

3.3.2. Knowledge Modeling

After the knowledge is acquired from different sources, the next step was organizing and structuring of knowledge. Knowledge modeling is the representation of information in the form of logic for the purpose of processing knowledge to simulate intelligence (Makfi, 2011).

Knowledge acquired through different knowledge acquisition techniques can be modeled with decision tree and hierarchical tree structure. Decision trees are produced by algorithms that identify various ways of splitting a data set into branch like segments. These segments form an inverted decision tree that originates with a root node at the top of the tree. The hierarchical tree diagram provides the analyst with an effective visual condensation of the clustering results. The hierarchical tree diagram is one of the most commonly used methods of determining the number of clusters. It is also useful in spotting outliers, as these will appear as one member clusters that are joined later in the clustering process. The numbers at the top and bottom of the hierarchical tree diagram represent equally spaced values of the criterion function. It gives a pictorial representation of the criterion function information (Chen et al., 2003).

For this study, hierarchical tree structure was used to represent knowledge modeling. Hierarchical tree structure can easily model concepts and clearly explains the concepts in the problem area. It models the knowledge in the hierarchical manner. This model starts from the main concept at the highest level of the hierarchy and other sub concepts that can affect or affected by the highest level concept put next to down ward in the hierarchy (Yemisrach, 2010).

The general structure of creating knowledge modeling contains input, knowledge model and output as shown below (Makfi, 2011).

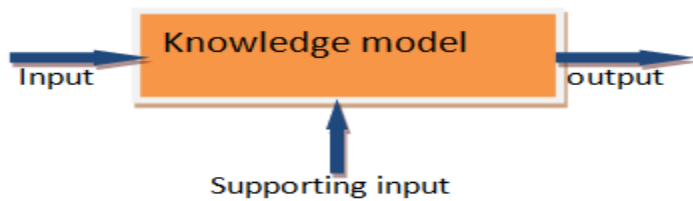


Figure 3. 3 knowledge modeling

The main inputs of the recommender system are students' information such as result/score, interest, special skill and health status which are provided by the student to the system and supporting inputs are knowledge elicited from domain experts, relevant documents and websites. Finally, the output will be recommendations for the student that enables them to make decision easily in their field of study selection process.

3.4. Hierarchical structure of field of study

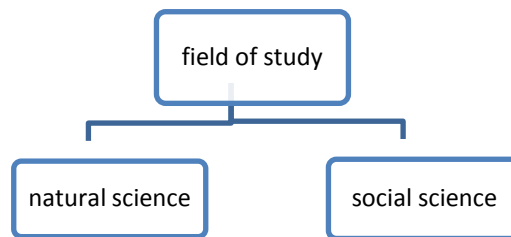


Figure 3. 4 Structure of field of study at preparatory school

As described by Ministry of Education of Ethiopia, students completed their preparatory school can be either of Natural science or social science students. Based on the new educational curriculum of Ethiopia, students are required to join higher education with related field of study. Social science students can join band 5 and band 6 while natural science students can only join band 1 to band 4.

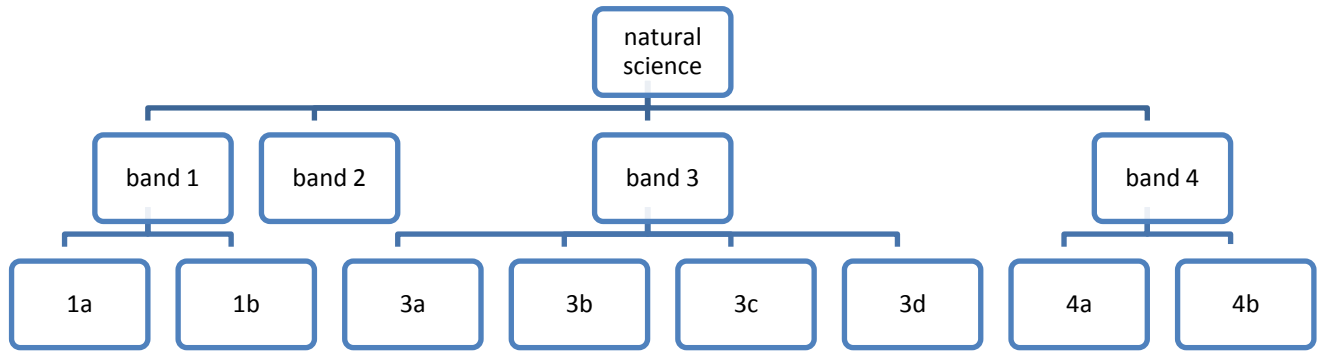


Figure 3. 5 General concept of Natural Science field of study

Keys:

Band1=Engineering and Technology, 1a= Engineering science, 1b= computer science

Band2= Natural and computational science,

Band3 Medicine and Health science, 3a= medicine, 3b= public health officer, 3c= other health science, 3d= dental science

Band4= Agriculture and Natural science, 4a= veterinary medicine, 4b= other agriculture and natural resources.

Natural science students have a range of fields of study to choose from if they fulfill the minimum requirement of higher education entrance examination of the year. Ministry of education grouped this field of studies in to bands and sub-bands that represent each field of study and sub-fields under each band. Natural science includes from band 1 to band 4; engineering and technology, natural and computational science, medicine and health science, and agricultural and natural science. Students are required to select their field of study in a descending order based on their interest and bands which represent field of study.

1. Engineering and Technology (Band 1)

Band 1 is the name given for all engineering and technology related field of study and grouped into sub-fields called engineering science and computer science. Each field of study contains departments/ schools as shown in the following table.

band	Engineering and Technology	
band 1	Engineering Science (1a)	Computer Science (1b)
	Civic, construction and transport engineering	Computer Science
	Environmental engineering	information system
	construction and building technology	information technology
	drafting and surveying technology	information science
	electrical and electronic engineering	management information system
	chemical and material engineering	computer engineering
	urban and regional planning	software engineering
	mechanical engineering and mechanics	Library science
	food technology and process engineering	
	sugar engineering	
	soil and water engineering and management	
	wood science and technology	
	water resource and environment engineering	
	water supply and environment engineering	
	biomedical engineering	
	irrigation and water resource engineering	
	architecture and surveying	
	agri-bioproass and bio-logical engineering	
	textile engineering and lather technology	

Table 3. 1 Field of studies under Engineering and technology

2. Natural and Computational Science (band 2)

Natural and computational science field of study is represented by band 2 and contains the following departments/schools.

- Biology
- Chemistry
- Mathematics
- Statistics
- Physics
- Earth science and geology
- Environmental science
- Health and physical education
- Meteorology
- Archeology
- Petroleum engineering
- Biotechnology
- Ecotourism biodiversity

3. Medicine and Health Science (band 3)

Medicine and health science field of study is categorized in to four sub fields as shown below.

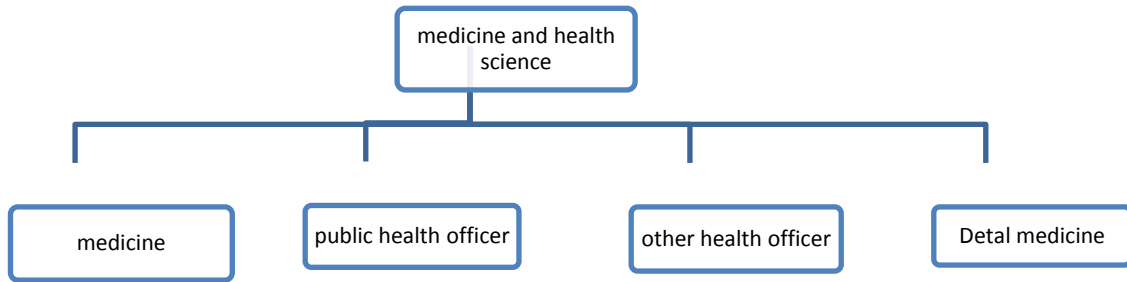


Figure 3. 6 Hierarchical structure of medicine and health science

Other health sciences field of study or school contains different field of studies/ departments as shown below.

- Medicine laboratory science
- Midwifery
- Nursing
- Psychiatry
- Pharmacy
- Environmental health science
- Health education and promotion
- Optometry/ physiotherapy
- Occupational health
- Radiography
- Dentistry
- Anesthesiology

4. Agriculture and Natural Recourse Science (band 4)

Agriculture and natural resource science field of study is grouped into two sub-bands; Veterinary medicine and other agricultural science as shown below.

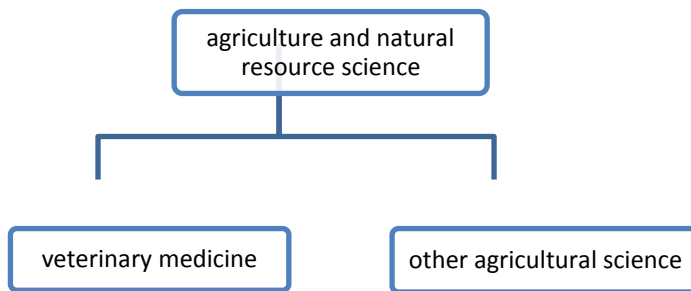


Figure 3. 7 Agriculture and Natural Science field of study

Other agricultural science contains different field of study as listed below.

- Range management and pastoral development, ecotourism
- Irrigation agronomy and engineering
- Horticulture, crop science
- Land and water resource management
- Land resource management and environmental protection
- Plant science and crop production
- Animal and wild life science, animal production
- Rural development, agricultural extension
- Cooperative and marketing cooperation
- Agri-business management
- Agricultural mechanization
- Forestry post harvest management
- Dry land agriculture
- Natural resource management
- Eco-tourism of wildlife management
- Soil and water management
- Agri-bioproass engineering
- Veterinary lab technology
- Disaster risk management and sustainable development
- Crop science
- Human nutrition

Student who complete their field of study with natural science can higher education with the above related field only.

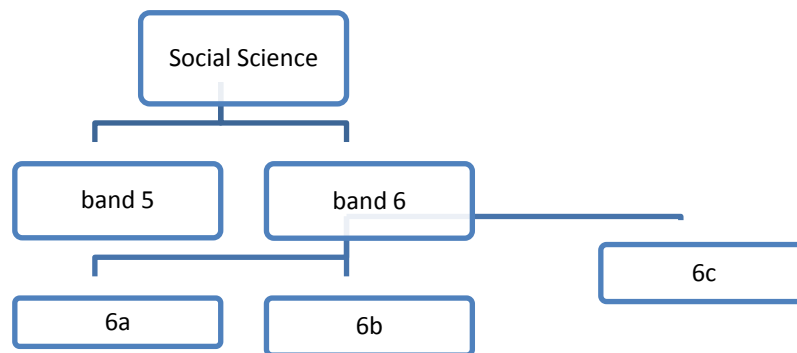


Figure 3. 8 Social Science field of study

Key:

Band 5= Business and Economics, Band 6= Social Science and Humanities, 6a = law, 6b= other social science and humanities, 6c= Theatrical, Fine Art and Music

Social science students are categorized under band 5 and band 6. Band 5 is all about business and economics field of study and band 6 contains law and other social science and humanities field of study which includes many fields.

5. Business and economics (band 5)

It is one part of social science and humanities field of study and contains the following departments and schools.

- Economics, finance and development economics
- Business management and administration
- Public administration
- Accounting and finance
- Banking and finance, banking and insurance
- Management, development management
- Marketing and sales management
- Tourism and hotel management
- Logistics and supplies management
- Business administration, management and information system
- Procurement and sales management
- Cooperative and marketing
- International trading and investment management
- Secretarial science and office management

6. Social Science and Humanities (band 6)

Social science and humanities field of study can be categorized into three as shown follows.

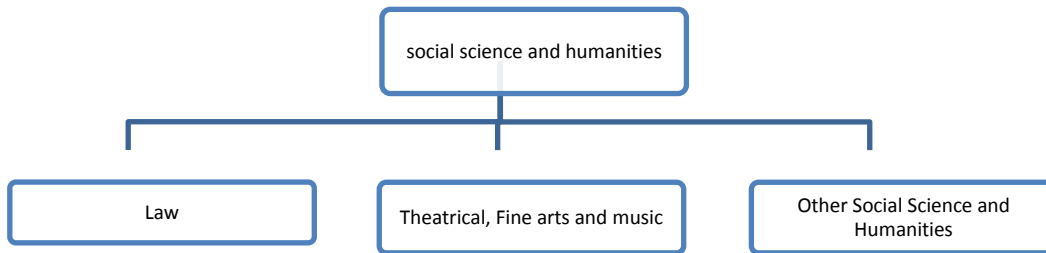


Figure 3. 9 Social science and humanities

Other social science and humanities field of study also includes many departments or schools as listed below.

- English language
- Arabic
- French
- Geography and environmental science
- Ethiopian, Amharic language literature
- Fine arts and design
- Population studies
- History and civics
- Journalism and communication
- Sociology and social work
- Anthropology
- Linguistics
- PSIR and ethics, civic and ethical education
- Gender and development studies
- Governance and development studies
- Teacher education and curriculum studies
- Educational planning and management
- Educational economics and management
- Adult education and community development
- Special needs education
- Psychology
- History and heritage management
- Archeology
- Philosophy
- Ethiopian language and literature
- Ethiopian sign language and deaf culture
- Ethiopian language and literature, Tigrigna
- Afan Oromo and literature
- Music, wind, string, piano

In general, students who fulfill the minimum entrance exam result or score requirement of the year should select their field of study based on their interest. Recommender system will assist them to choose their field of study correctly based on their interest by providing recommendation among many fields of studies.

CHAPTER FOUR

DESIGNING AND IMPLEMENTATION OF THE PROTOTYPE

The design and implementation part of this section involves the actual development of a workable case based recommender system for field of study selection process in the higher education institution of Ethiopia.

Relevant cases are collected from domain experts and students. It was done by organizing relevant attributes to be filled by domain experts and students. (Main attributes are discussed in the next section). Having all the relevant cases and knowledge from the domain expert, students and different relevant documents, the next task is coding the knowledge into computer using appropriate and efficient knowledge representation methods and tools. For this research, JCOLIBRI 1.1 is used to design and implement the prototype and construct the case structure of recommender system.

The main algorithm used in this research is Nearest Neighbor retrieval algorithm. This is because JCOLIBRI uses nearest neighbor algorithm for retrieval task. Nearest neighbor algorithm retrieves the case which is nearest to the user's query by measuring its similarity with the cases. Given a collection of cases and query point in an m -dimensional metric space, find the new case that is closest to the query point. Similar queries are performed by taking a given complex object, approximating it with a high dimensional vector to obtain the query point, and determining the data point closest to it in the underlining feature space.

The development of recommender system application involves a number of steps, such as collecting case and background knowledge from different sources, modeling a suitable case representation that are used for decision making, defining an accurate similarity measure, implementing retrieval functionality, and implementing user interfaces. The main feature of JCOLIBRI programming tool is used to deliver the actual prototype of the system in this research. JCOLIBRI has been constructed as core modules to offer the basic functionality for

developing case based recommender system application. Implementing a recommender system application remains a time consuming software engineering process and requires a lot of specific programming experience beyond pure programming skills (Stahl & Roth-Berghofe, 2008). Using JCOLIBRI minimizes the effort to develop recommender system by using other programming languages such as java.

4.1. Designing the Architecture of Case Based Recommender System

Figure 4.1 show that the framework of case based recommender system for student placement in the higher education. To develop recommender system, the researcher collects the important knowledge from relevant documents, students and domain experts. In designing field of study recommender system, the researcher collects the relevant attributes and cases from domain experts and students and the required knowledge was represented. Thus building of case based recommender system was started by collecting the previously solved cases about field of study process from students who succeed in their correct choice. The collected raw cases needs further processing to get the required case base on the problem stated in order to represent case structure and important attributes. After processing of cases and having the selected attributes, assigning weight and important parameters for each attribute was the next performed step. This is because since all attributes are not equally important to recommend field of study for the students.

Once the case based recommender system is developed, users/students can use the system easily to choice their field of study based on the recommendation given by interacting with the system in order to retrieve the best cases that can match with their query. When the student enter their query/case description through the user interface window, the system searches the best matching cases from the case base and retain the possible solution. If there is exact matching between the query and cases in the case base, the system recommends the most matched field of study for the student based on the given query that is used to state once interest. If the similarity between query and case is approximate, the proposed solution needs modification (adoption of solution) to fit the current problem described. At the end, the best

modified solution should be stored into the case base for future use. The case base updates incrementally when the system learns from new case used by the student.

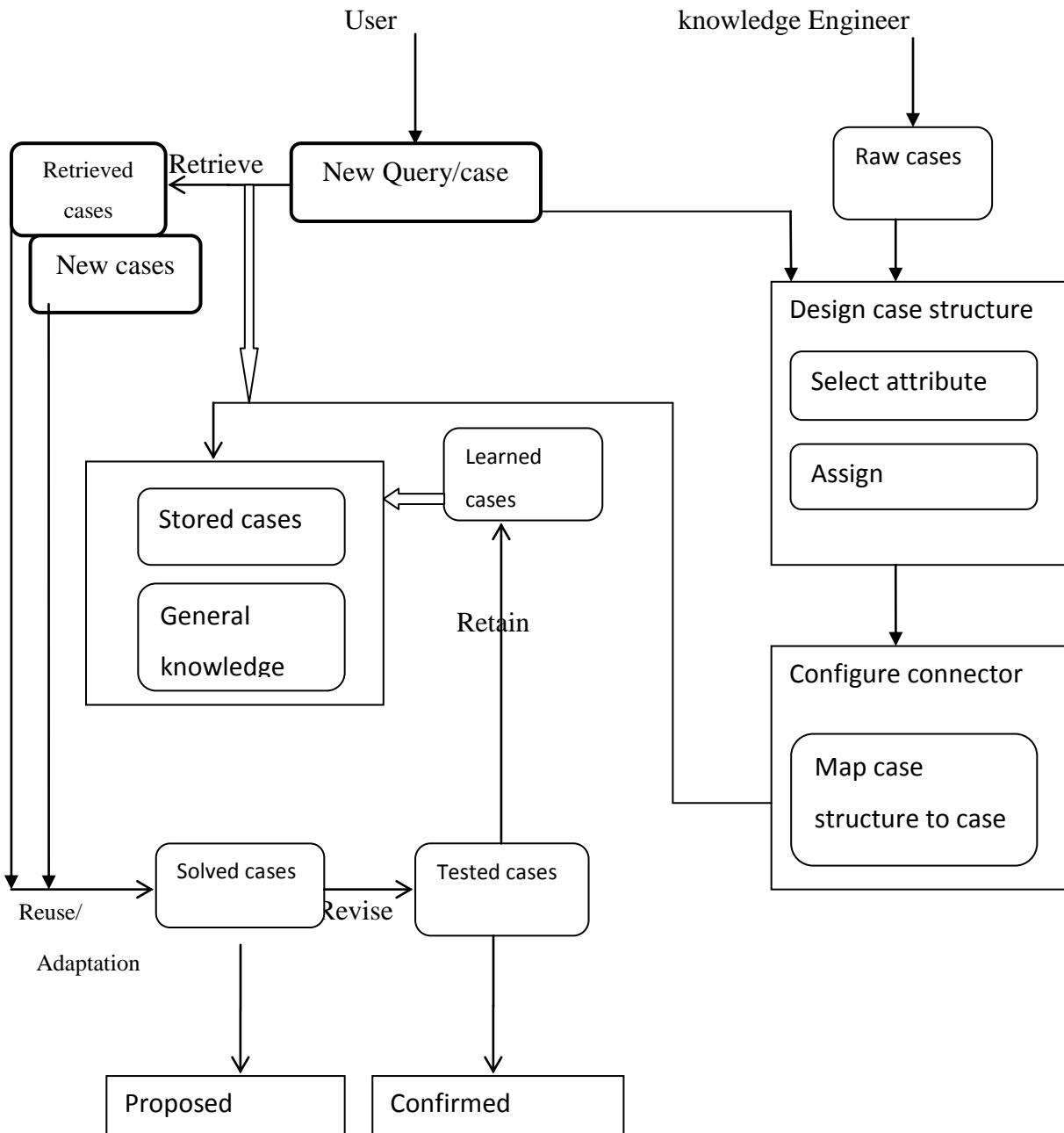


Figure 4. 1 Case Based Recommender System Architecture

4.2. Case Based Recommender System for field of study selection

To develop case based recommender system, JCOLIBRI case based reasoning framework is used. JCOLIBRI has been constructed as core modules to offer the basic functionality for developing case base reasoning as well as case based recommender system.

Developing a new case base recommender system is made by writing few Java classes that extend classes of the framework and configure some XML files.

To start the JCOLIBRI graphical user interface (GUI) application tool, launch the main window by clicking on **JColibriGUI.bat** file and it becomes ready to use as shown below in figure 4.2

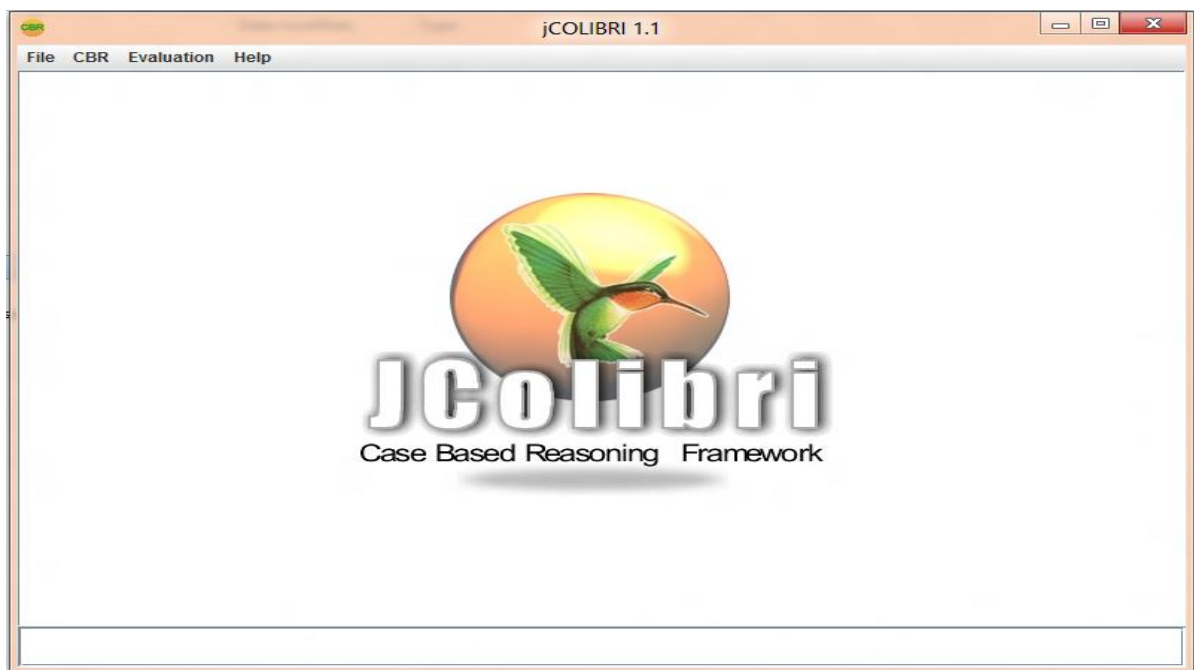


Figure 4. 2 JCOLIBRI main windows

In this study, the development of case based recommender system for field of study selection process can be divided in the following sub processes which enable to achieve the objectives of the research.

4.2.1. Building the Case Base

As stated in the objective of this study, one of specific objectives was building case base comprising of field of study selection process cases to provide the most important similar cases

to support students in the selection of their field of study. As a result of this, the researcher collected the real case from successful students. The acquired cases are used to build case based recommender system for field of study selection that is important to assist students in their field of study selection process. Cases are stored in the case base as a text file in structured format after the researcher has analyzed and interpreted the case.

The case base is presented as a plain text comprising of N columns representing case attributes ($A_1, A_2, A_3, \dots, A_N$) and each M rows representing individual cases C ($\{C_1, C_2, C_3, \dots, C_M\}$) each attribute has a sequence of possible values associated to each column attribute $A = \{V_1, V_2, V_3, \dots, V_k\}$.

The case base consists of a set of cases that represents knowledge about field of study selection processed by the successful student. The researcher tried to collect all the cases by selecting successful students from Debre Markos University due to right selection of field of study. These students are selected purposefully from each field of study who are joined any of their three choices.

4.2.2. Case Representation

The case representation has been formulated in the way that easily represented in JCOLIBRI. Designing of such case structure helps to define the features available in the cases and used to measure the similarity between existing cases and the new case (query). The general application of this research is to retrieve similar cases to the query from the case base that can guide students, solving problems of confusion and transforming a recommendation in their field of study selection process. Case base were structured to make the retrieval process efficient. This is done through case indexing process in the JCOLIBRI programming tool. Indexing refers to assigning index to the case for retrieval by comparing the existing case and the query given by the user.

4.2.2.1. Attribute selection

After discussing with domain experts, reviewing of selected attributes have been made to identify best attributes for the system that can recommend field of study. These attributes are

collected from 32 domain experts of Debre Markos University from each department head. The researcher prepared interview question for domain experts or department heads to collect important attributes.

Interview questions contain the following attributes to be selected by domain experts. Age, sex, Background interest, Family interest, Result , Relate subject result in preparatory class, Preparatory class attend, Entrance exam result, Health condition, language dependency, duration, past experience /worked on related field , need of nearby university to family, economic back ground, job opportunity, pre-request subject taken at preparatory level. These attributes are proposed by the researcher by considering the real situation. In addition to these, the researcher tried to find additional attributes from experts but unfortunately there are no additional attributes found from domain experts.

The attributes selected by domain experts and their frequency are listed below.

Name of attributes	Frequency
Preparatory class attend	6
Special skill	3
Interest 1, 2, & 3	6
Background interest	6
Family interest	6
Preparatory best result	3
Entrance exam result	6
Health status	6
Language dependency	2
Duration	4
Past experience	3

Table 4. 1 Selected attributes by domain experts

As shown in the above table 4.1, from the given sixteen options, domain experts selected thirteen attributes but the remaining attributes were not selected by any expert. The researcher used the thirteen attributes as the main criteria to assign students in different field of study. Preparatory classes attend, interests, background interest, entrance exam result, and health status attributes have been selected by six experts. This shows that they are important

attribute to assign students in different field of study. Special skill, preparatory class best result and past experience attributes have been selected by three experts. Therefore they are important criteria to assign students to field of study based on their special skill, preparatory class best result and their past experience. Family interest and duration attributes are selected by four experts while Language dependency attribute has been selected by only two. All these selected attributes are then considered important to assign students to field of study. The selected attributes are described as follows.

Preparatory class attend: - this attribute contains only two options, social science and natural science. It indicates the background of students` in preparatory level.

Special skill: - this shows the previous skill of the student that enables them to select appropriate field of study based on their skill. For instance a student may have computer skill to join Information technology or computer science department; technical drawing skill to join engineering; social interaction skill to join psychology or sociology; art to join theatrical art and soon.

Interest 1, 2, & 3: - the first three best choice of field of study selected by students. These are most important attributes to assign students based on their interest.

Background interest: - This is the interest of the student in their preparatory and elementary level. Students may want to be a doctor or an engineer, or pilot and soon from their lower level. Therefore, students are required to set their background interest to this attribute.

Family interest: - this is the interest of student`s family that advise their students to select best field of study based on their interest. For instance student may want medicine where as their family may need engineering. The system will try to measure similarity of interest based on given case.

Preparatory best result: - this attribute indicates subjects in which the student had maximum score in their preparatory class. For instance if a student has maximum Score with Information technology, he/she is recommended to join computer science related field.

Entrance exam result: - this shows the entrance exam result of the student that might be required to assign students in different filed. For instance students who have 600 from 700 in entrance exam result may recommend to join medicine or engineering since his/her result show his/her academic performance.

Health status: - this indicates the health condition of the student. For instance student who is blind or deaf may not be recommended to join sport science rather they are recommended to join history or language departments.

Language dependency: - this is to indicate the language problem to assign students. For instance if a student understand only Amharic language, he/she is not recommended to join AfanOromo department or Tigrigna department.

Duration: - this shows the interest of students to attend their university level and spent in the campus. For instance some student may need to graduate not more than 3 years due to many factors such as need to help their family. If a student needs to graduate with in 3 year, they are recommended to join fields that have short duration.

Past experience: - if a student has experience on something, they are recommended to join related field. For instance if a student has diploma in Information Technology and worked few years with Information technology, she/he is recommended to join Information technology.

4.2.2.2. Description and Weight of the Selected Attributes

The case in this research consists of thirteen descriptions/attributes that served to contain descriptions of the problem which is used to make decision by the system and two solutions attributes which holds solution for the recommendations. The following table shows the description of selected cases with their value and Weight.

Significant Attributes			
Attribute Name	Data Type	Weight	Local Similarity
Preparatory class attend	String	1.0	Max string
Special skill	String	0.9	Max string
Interest1	String	1.0	Max string
Interest2	String	0.95	Max string
Interest3	String	0.9	Max string
Background interest	String	1.0	Max string
Family interest	String	1.0	Max string
Preparatory best result	String	0.9	Max string
Entrance exam result	integer	0.95	Equal
Health status	String	1.0	Max string
Language dependency	String	0.8	Max string
Duration	String	0.9	Max string
Past experience	String	0.9	Max string
Solution Attributes			
Recommended field	String	1.0	Equal
Band	String	1.0	Equal

Table 4. 2 Descriptions and Weight of the Selected Attributes

The above table shows the general description of attributes consisting of attribute name, data type, weight and local similarity. The most significant attributes to the problem domain such as preparatory class attend; family interest, interest1, health status and background interest have the highest weight value of 1.0. These attributes are the most relevant to the students to select field of study based on their interest. Next to these, attributes like interest2, entrance exam result, interest3, special skill, preparatory best result, duration, language dependency and past experience have the weight value of 0.95, 0.95, 0.9, 0.9, 0.9, 0.9, 0.9, and 0.85 respectively. The assignments of weights to each attribute indicates that attributes having high weight is the most relevant to the user (students) in the distribution process of field of study selection. The weight of each attribute has been assigned its value by domain experts at the time of attribute selection. The local similarity of most attributes is maximum string. This is due to the similarity between query and cases can be calculated with maximum string length. Few attributes such as

entrance exam result have equal similarity weight since it needs exact match of cases and new query.

After identifying relevant attributes of the case, the next task is definition of appropriate similarity measure in JCOLIBRI. JCOLIBRI follows both local and global similarity measures. Local similarity measure divides the similarity definition into a set of local similarity of each attribute where as global similarity calculates the final similarity measure.

Different types of local and global similarities are used in this research.

Local similarities

- ✓ **Equal:** - The input query and cases in the case base must match to get the result; if there is no match between input query and cases, matching will fail.
- ✓ **Interval:** - Exact match is not required in this similarity. When it is assigned to the attributes, JCOLIBRI reminds this interval value in the searching from the case to get the similarity.

Global similarity

- ✓ **Average:** - it is a type of global similarity that considers the average of all attribute similarity values.

4.2.3. Managing Case structure

One of the most important features of JCOLIBRI is managing or defining case structure of the system easily by case structure window. The selected attributes were added to the description window of case structure and their properties are assigned for each attribute to the right side of the window. After this, the program generates a java code automatically and saves in XML format. As shown in the following figure 4.3, data type, weight and local similarity of the selected attribute can be configured on the right side of the case structure window. The window is divided into two parts. The left side contains the structure of the attributes as a tree. We can generate description as well as solution cases by clicking “add simple” button and

remove unwanted attributes by clicking on “remove” button. The right side contains the property values of the attribute which can be configured as the property of cases.

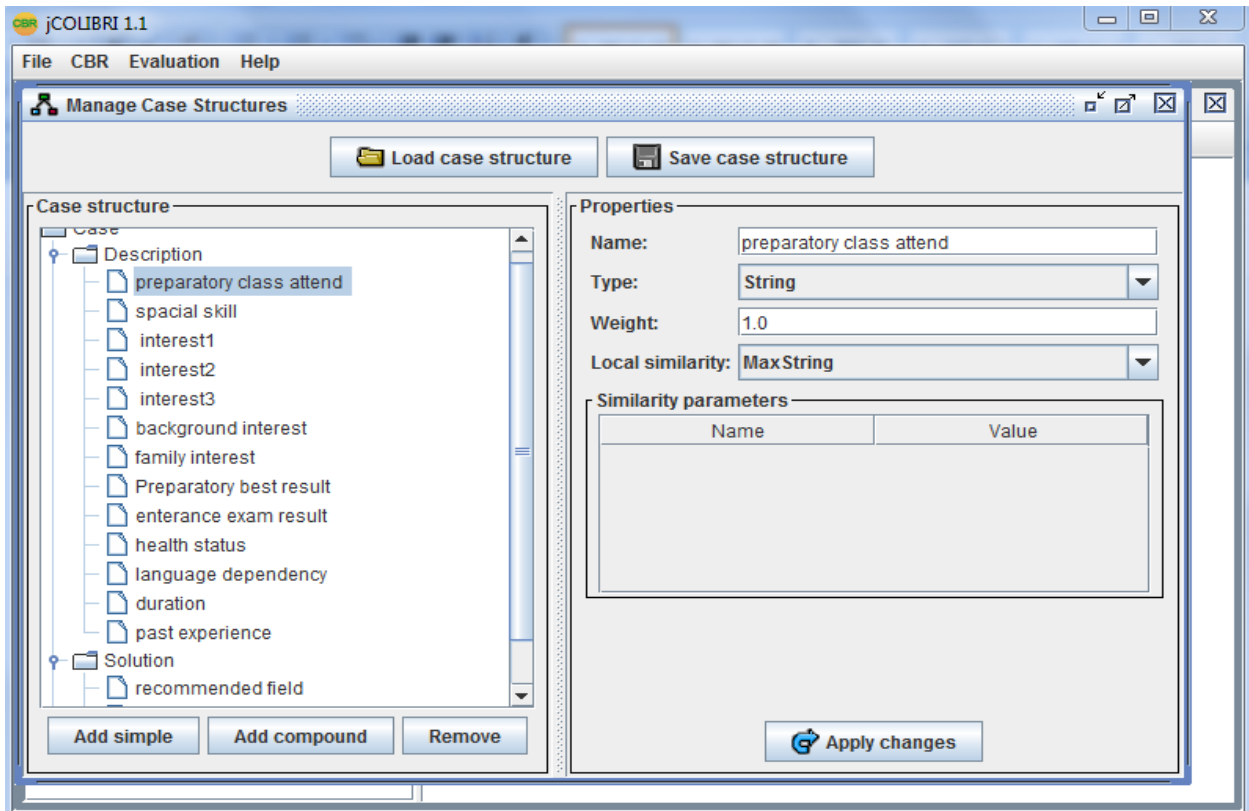


Figure 4. 3 Managing Case Structure

4.2.4. Managing Connectors

Once case structures are configured in JCOLIBRI, recommender systems must access the stored cases from case base. JCOLIBRI supports both SQL database and plain text file to store its cases base.

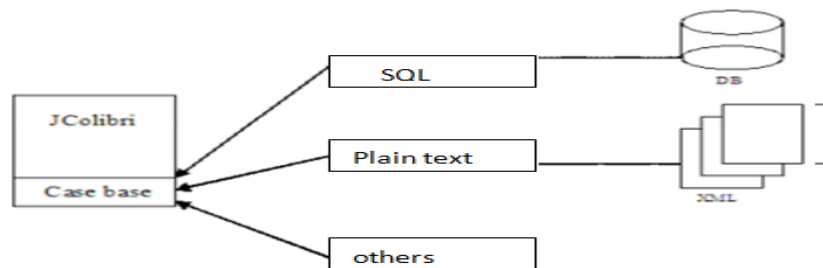


Figure 4. 4 JCOLIBRI case base schemas

In this research plain text connector is used as a case base storage. The connector maps the case structure to its column from plain text file which is saved in **.txt** file format and later saved as XML file like that of case structure.

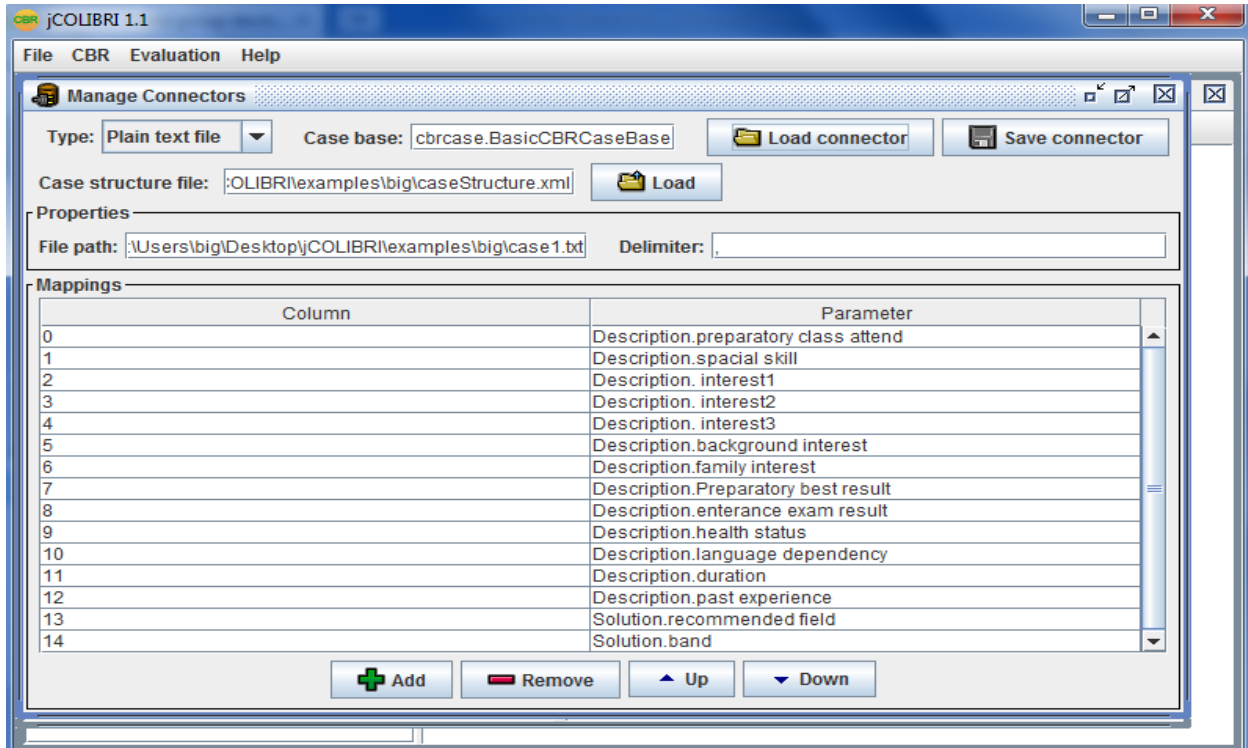


Figure 4. 5 Managing Connectors

One of the most important tasks in managing connectors is specifying the correct path of case structure and file path. The case structure path is used to access and match attributes from case structure and file path is used to specify the **.txt** file that contains the case base. Delimiter of this connector uses comma (,) to separate value of each attribute in the case.

4.2.5. Managing Tasks/ Methods

JCOLIBRI is organized into packages. These packages can perform and execute tasks and methods decomposition process. For the development of case base recommender system prototype, the researcher used core package task. These core packages are PreCycle, main CBR cycle and PostCycle. The detail of each tasks and methods can be discussed separately as follows.

4.2.5.1. Managing Tasks

After configuring the connector and case structure, the next task is selecting tasks and methods of application. Different core packages are available, which are the most important packets in JCOLIBRI. The main components of core packets are discussed as follows.

- **PreCycle:** - This part of the task retrieves data or cases from case base before execution of the main cycles.
- **MainCBRCycle:** - it retrieves the most similar tasks and describes the typical cycle task at the highest level and obtains the query. Even if it is case based reasoning cycle part, it is also used to retrieve the most similar task for recommendation system from stored cases; reuse previously stored cases as knowledge to solve the problem; revise the proposed solution and lastly retain the experience. In this cycle, there are other subtasks and most of the tasks requires specific path of the case structure. The case retrieval phase of CBR cycle involves finding similar cases with the query using selected algorithm. After searching and selecting most similar cases with the new problem, the solution of problems will be displayed to the user. This task has sub tasks again, select working cases task which selects working cases from case base and stores into current case base; compute similarity tasks which is used to compute the similarity of cases between existing cases and new query; select best case which returns best cases from case base with high degree of similarity with new case.

Case Similarity, Matching and Ranking

The main goal of case base recommender system is to retrieve the best similar cases to the query from case base and selecting the nearest similar case. Selecting the best similar case is usually performed by means of some evaluation heuristic functions or distances, which are possibly domain dependent. JCOLIBRI uses the nearest neighbor algorithm as case retrieval technique. Nearest neighbor algorithm used to measure the similarity between the existing cases and the new case/queries, and then return the search results in their ranked order. The local similarity function measures the similarity between each and every simple attribute values in the case base with new cases queries. The similarity score will be assigned based on the matching weighted sum features from those simple attributes.

The average score of each attribute between existing case and query are computed and the similarity between stored case and the query result is assigned to the object. Finally, the maximum degree of similarity among retrieved cases is displayed in their ranked order.

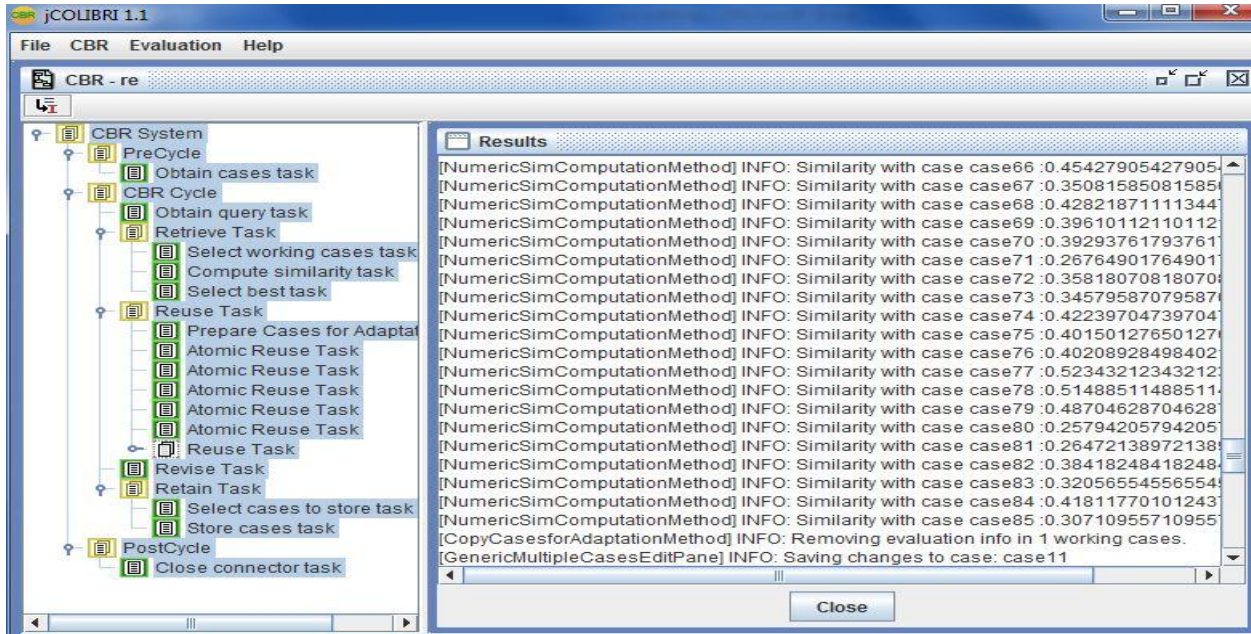


Figure 4. 6 case similarities between case base and query

Reuse/ Adaptation: During case retrieval, once one or more very similar case is identified, the solutions are selected for this particular problem to meet the requirement of new solution. This reuse stage generates the proposed solution for the problem. There are situations where cases are not similar with the new case. At this time, this new case can be stored in the case base and will be reused by other students for the next time. The system can learn at every entry of new case and new users adopt this knowledge for field of study selection process.

Revise task: This stage is the evaluation stage about the selected solution in the reuse phase. After selecting the most similar cases from search result, the solution for problem should be confirmed and validated before the solution is stored as a case for future use.

Retain task: after having confirmation in revision phase, the problem together with its solution will be stored in case base. These tasks are after validation and confirmation of retrieved case solutions at reuse and revise phases.

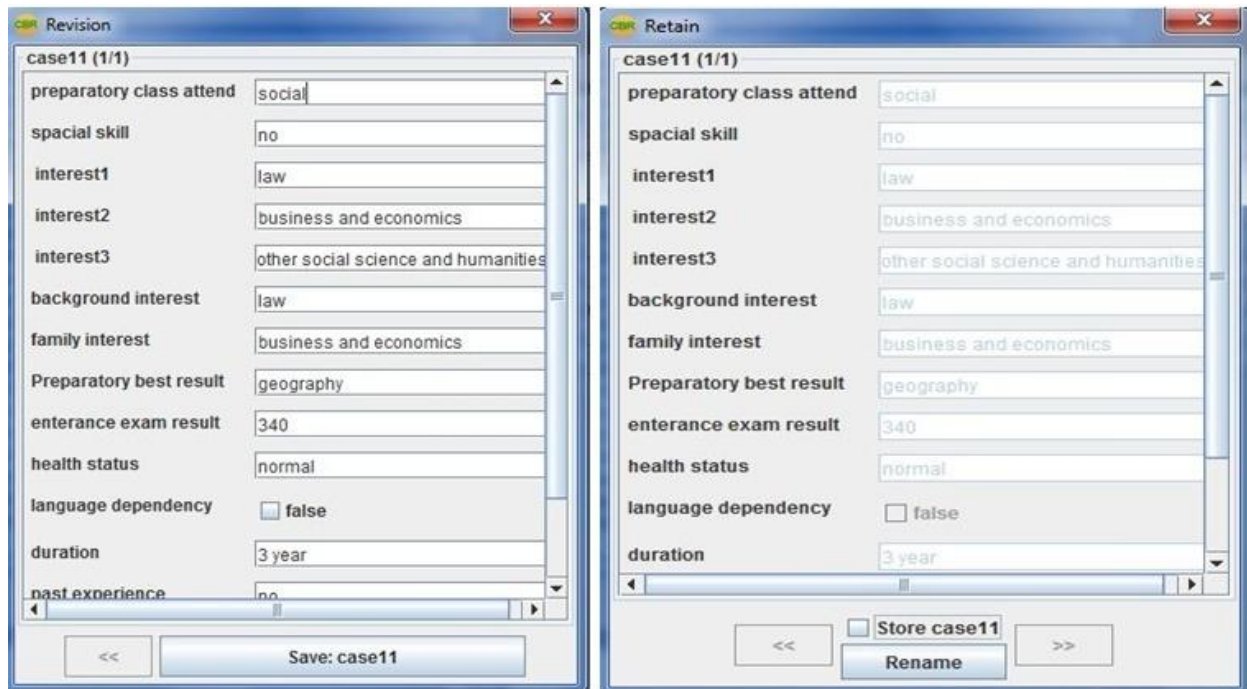


Figure 4. 7 Revision and retain tasks

- **Post Cycle:** - this cycle contains tasks that executes after main CBR cycle. Case retrieve begins with case description and ends with best matching of the case with the new case. The subtask of postCycle task called close connector task will close the connector and save the case base.

4.2.5.2. *Managing Methods*

The method packages store classes that resolve the task. These classes can resolve the CBR cycle using programming or using graphical user interface (GUI). All tasks in JCOLIBRI should have their own methods to be assigned in order to achieve its recommendation goal.

LoadCaseBaseMethod: This method returns the whole available cases from the case base to designer and use connector as parameter to retrieve case base.

ConfigureQueryMethod: This method obtains and configures the query. It displays the graphical user interface window by receiving case structures as an input to request query and to receive cases from the case base.

SelectAllCaseMethod: Selects working cases from case base and store them into current context. It allows displaying all the available cases from the case base to the result window.

NumericSumComputationalMethod: Computes similarity tasks between cases and query.

ManualRevisonMethod: Manual revision method enables users to modify cases in the query window.

SelectSomeMethod: select best of found cases. It returns most similar value of the top best selected case.

Tasks in JCOLIBRI can be solved with different methods as listed above. Choosing the most appropriate method for the task is the role of researcher in the designing of case base recommender system. For this research, only few of them are selected and discussed which are appropriate for recommendation system. Figure 4.6 shows the main window of cycle of JCOLIBRI tasks and methods. As shown below, pre cycle, main CBR cycle and post cycle are on the left side of the window. When the designer selects any task from these cycles, the configuration method windows displayed on the right side and appropriate inputs can be selected according to the situation. These inputs are parameters for new instances.

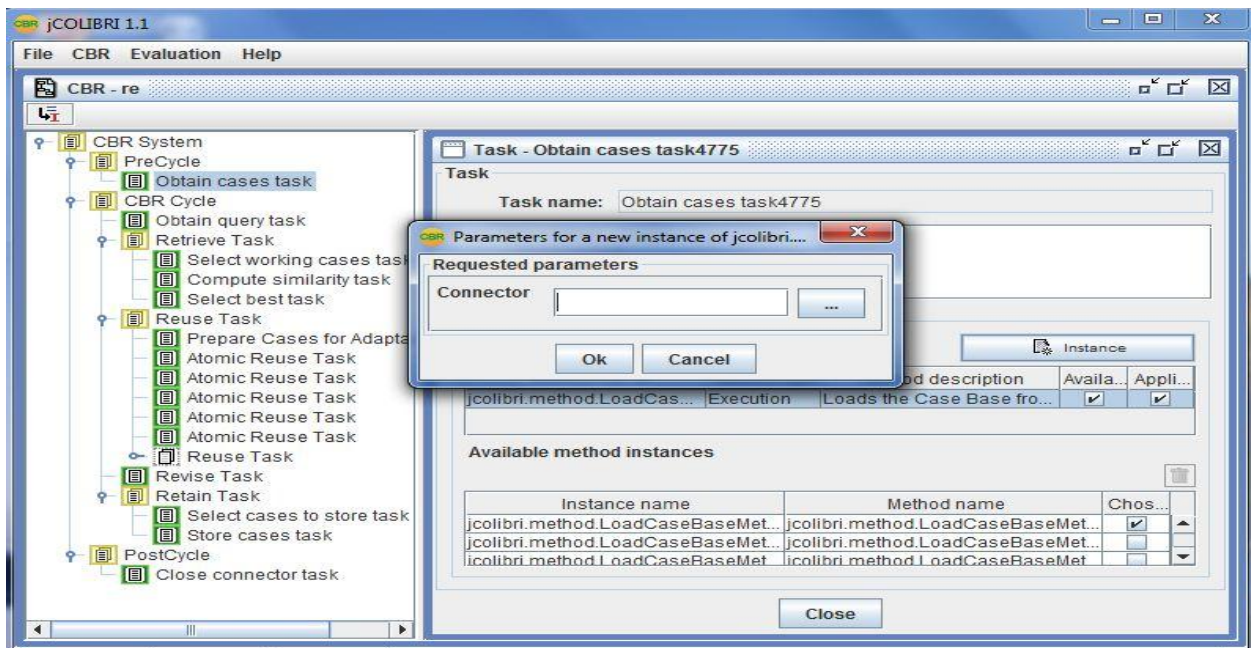


Figure 4. 8 Managing Tasks/Methods

4.2.6. Deploy the case base recommender system

After defining and configuring all the necessary steps required designing case base recommender system in JCOLIBRI, testing the recommender system application is the next step as shown in figure 4.7.

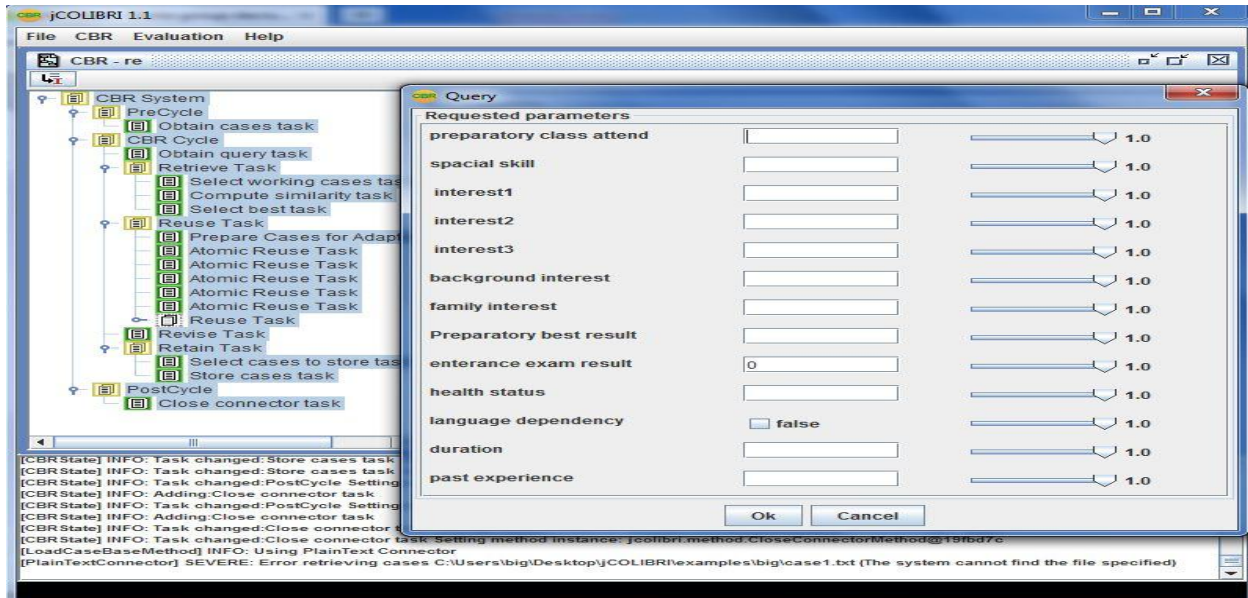


Figure 4. 9 Case entry windows

Students are required to enter the query on the space provided according to the instruction. In the “preparatory class attend” box, students are required to enter their background level; either social science or natural science. Students who attend social science should select social science field of study and the same is true for natural science. Students may have special skill and attitude toward some field of study. If they have skills that enable them to join related field, they required to enter that skill in the “spacial skill” box. Interest1, interest2 and interest3 are interest of students they want to join like engineering science, medicine, law, etc. They are required to enter their interest accordingly based on their interest. Background interest and family interest needs to enter background interest of the student when they are in lower grade level such as to be a doctor and engineer and their family interest about the student. In the “preparatory best result” box, students are required to enter the subject which scored best result in the preparatory class where as in the “entrance exam result” box students are required to enter their average result of higher education entrance examination. In the “health

status” box, students are required to enter their health status either normal, blind, deaf, or physically hand cape. This is important to assign students to field of study which best matches their health condition. Students may have language problem to assign field of study. For instance, if a student doesn`t know Amharic language, she/he is not recommended to join Amharic department. Students are required to enter their language status with yes or no options in the “language dependency” box. In the “duration” box students are required to enter three year or more than three year. This is due to students may need to graduate with in three year or may not consider year level. In the “past experience” box, students required to enter their past experience if they have or ‘no’ if they don`t have any experience.

CHAPTER FIVE

TESTING AND PERFORMANCE EVALUATION OF THE PROTOTYPE

5.1. Experimental Setting

Construction of case base is the first and most important step in developing case base recommender system application. After selecting thirteen important attributes, constructing cases is the most challenging task in developing case base recommender system. As discussed in the fourth chapter, these attributes are selected by consulting experts from different areas. The researcher tried to organize interview questions to get appropriate attributes which can help students to make the right decision when they select field of study. Each field of study requires students to have certain attributes in order to succeed in the field. For instance engineering science may need drawing skill, theatrical art may need art skill and soon. As replied by all experts, preparatory class attend, the three interest of students (interest1, interest2, interset3), background interest of the student, family interest, preparatory best result, entrance exam result, health status and past experience attributes are the most important criteria to be considered for every field of study selection. The remaining attributes such as special skill, language dependency and duration attributes are required by few field of study as the main criteria for students to succeed. For instance students who are from Amharic language speaker may have language problem to succeed in other language dependency field of study.

After selecting these important attributes, the next task was constructing the case base. These cases are organized based on important attributes (Appendix B). The researcher organized these cases in excel format and collected all the important data from relevant students.

The cases consist of various degree of complexity in terms of size of cases and attributes, the amount of parameters, weight and type of constructs. For this research a total of 105 cases are used to build the case base and to test the prototype. These cases are representatives of most

fields of studies and related criteria to join these fields. Therefore, these cases are important for students for their decision making when they select field of study.

5.2. Testing the main Cycles and Evaluating the Performance of the System

To check the validity and performance of the recommender system to domain experts, the functionality of case base recommender system main cycles and effectiveness of the prototype should be tested with selected cases. The effectiveness of the prototype is measured with recall and precision using test cases.

In addition to that, the performance of the system was evaluated from the users' side with users' acceptance testing. With users' acceptance testing, potential users' of the system rate the applicability of the system in their field of selection process.

5.2.1. Evaluation of the Retrieval and Reuse Process

Retrieval of previously stored cases from the case base to solve new problem by measuring the similarity of stored case and new query is the first step in JCOLIBRI for case base recommender system application. Retrieval of similar cases from existing cases to the new case is followed by the reuse of similar solutions with solutions of previously solved problem. Since the implementation tool JCOLIBRI uses nearest neighbor retrieval algorithm, retrieval of cases are performed using this algorithm. Nearest-neighbor retrieval technique is used to measure similarity between the source case and the case which we are searching. The nearest neighbor algorithm measures the similarity of stored cases with a new input case, based on matching a weighted sum of features. When a new case doesn't exactly match with old cases, then this algorithm will return nearest match from case base. But the retrieval time of this algorithm increases linearly as the case in the case base increases (Lang & Lau, 2002).

The nearest neighbor algorithm can be represented in the following equation (Watson, 1994).

$$NN(I, R) = \frac{\sum_{i=0}^n w_i \times \text{sim}(f_i^I, f_i^R)}{\sum_{i=0}^n w_i}$$

Where:

w is the importance weighing of an attribute, I is the target case, R is source case i is individual attributes from 1 to n , sim is the local similarity function, and f_i^I and f_i^R are the values for attribute i in the input case (I) and case in the case base (R) respectively and n is the number of attributes in the case base.

Similar cases to the new cases are retrieved with appropriate ranking order during retrieval process. The user of the system (students) can use the recommended case which is retrieved based on the solution cases in a way that can fit to the new query. The process of retrieval and reuse of cases is successfully implemented in case base recommender system application as shown in the following figure 5.1.

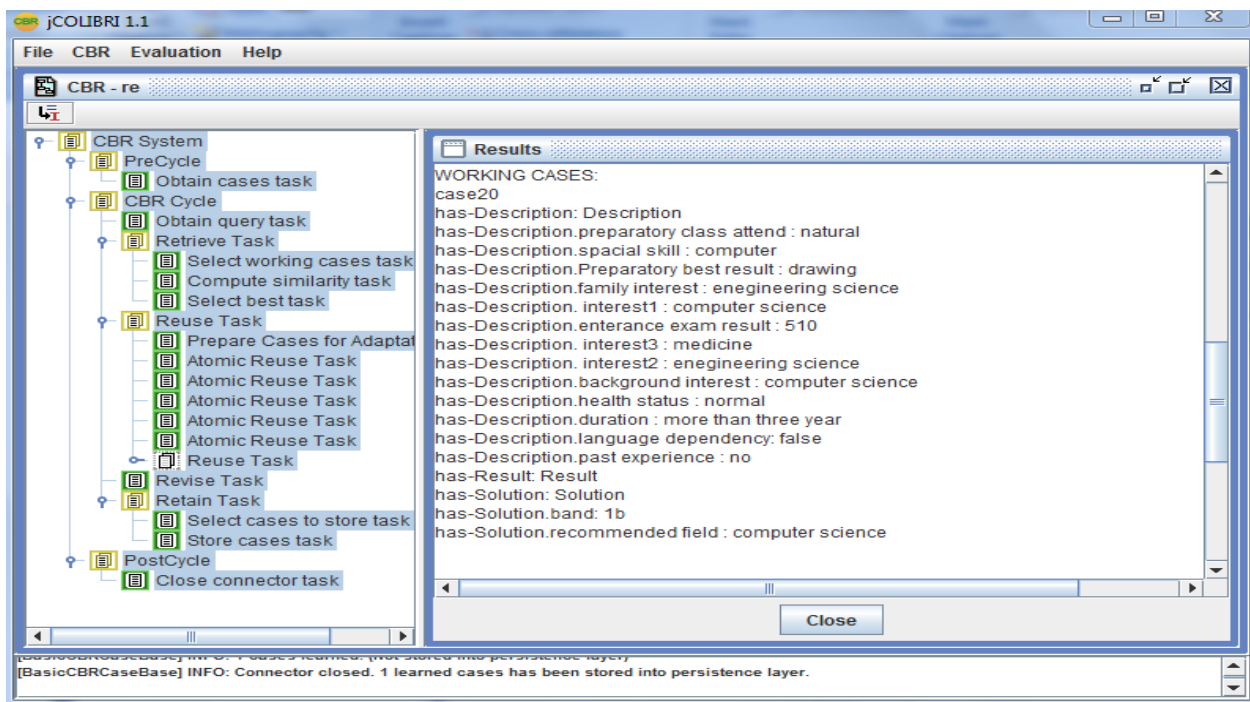


Figure 5. 1 Retrieval case

As shown in the above table, the system calculated the similarity of new case and existing cases and displayed the most similar cases to the new query. This is the recommended solution to the student based on the given case.

Precision and recall are the commonly used measures of performance of the retrieval process. Precision is the proportion of search results that are relevant to the query and recall is the ability of the retrieval system to retrieve all relevant cases to a given new query from the cases base.

$$\text{Precision} = \frac{\text{number of relevant cases retrieved}}{\text{Total number of cases retrieved}}$$

$$\text{Recall} = \frac{\text{number of relevant cases retrieved}}{\text{Number of relevant cases in the case base}}$$

To start the evaluation process, for each selected test cases the relevant field of study selection cases from the stored case base should be identified. These selected test cases were given to domain experts to assign relevant cases to each test query. Domain experts use the value of assigned field (recommendation part) attributes of the field of study selection cases as the main input to assign the relevant case to the test case as shown in the following table 5.1.

Test Case	Relevant cases from the case base
Case10	Case57, case62, case34, case16, case5, case38, case50, case69
Case13	Case26, case40, case67, case70, case18, case33
Case70	Case63, case44, case30, case19, case71, case67, case9
Case77	Case89, case99, case80, case102, case90, case97
Case64	Case21, case22, case13, case66, case39, case47, case20
Case28	Case3, case32, case55, case52, case19, case31, case45, case6, case11

Table 5. 1 Relevant Cases Assigned by the Domain Expert for Sample Test Cases

After the identification of the relevant cases to the test cases by the domain expert, precision and recall values were calculated with threshold interval to test the prototype using the test case query to know the performance of the system.

There is no standard threshold for degree of similarity that has been used for retrieving relevant cases. That is why different researchers use different case similarity threshold to measure the performance of their system.

After similarity computation, the first 6 best cases are retrieved assuming that the first 13 cases are selected as a case base for experimentation as shown in the following table (table5.2).

Test Cases	Recall	Precision
Case10	0.75	0.61
Case13	1.0	0.46
Case70	0.85	0.54
Case77	1.0	0.46
Case64	0.85	0.54
Case28	0.66	0.69

Table 5. 2 Performance Measurement of case base recommender system using Precision and Recall

As shown in table 5.2 above, recall of each test case can be calculated by dividing the number of relevant retrieved cases with total relevant cases. For instance, for case10 six relevant cases are retrieved out of the total eight relevant cases in the case base. When we calculate its recall value, the result is 0.75 (75%). All the other test cases can be calculated in such a way. When we come to the precision of the system, it is calculated as the number of relevant cases retrieved divided by total retrieved cases. For instance, case10 contains 8 relevant cases and total of 13 retrieved cases. Therefore the precision value of case10 is 0.61 (61%). The precision value of all test cases can be calculated in such a way. The average recall and precision value of the system is 85% and 55% respectively. Recall in information retrieval is the fraction of the documents (case) that are relevant to the query that are successfully retrieved. It is the ability of a retrieval system to obtain all or most of the relevant documents in the collection (Clarke & Willet, 1997). The higher recall value shows that the system obtains most of relevant cases from the case base. Therefore, this recommender system can retrieve relevant cases that enable students to make decision easily in field of study selection process.

On the other hand, the system retrieved relevant cases to the system with 55% precision. The precision value of the system is not as expected by the researcher due to few number of cases used. As the number of cases increased, the precision value of the system will also increase and better performance will scored in retrieving relevant cases.

5.2.2. Case Revision and Solution Adaptation Testing

The purpose of testing adaptation of solutions is to evaluate the systems capability to reuse existing cases from the case base. The system loads case base at the PreCycle stage of JCOLIBRI

framework and then selects working cases from the case base. These working cases always stored in to current context at the retrieval stage. The next stage is reusing the cases that are stored in the working memory and this reuse stage can be used by the next user of the system at another time. The adaptation process of case base recommender system is successful as the case features of the previous and new case have similar or less contradiction attribute values. Adaptation process will not be performed as the attribute values of the previous and new cases have more dissimilarity or totally different from the previous cases. In such cases, the adaptation process has to be edited and performed manually by a human domain expert in the revision stage as shown in the following figure 5.2.

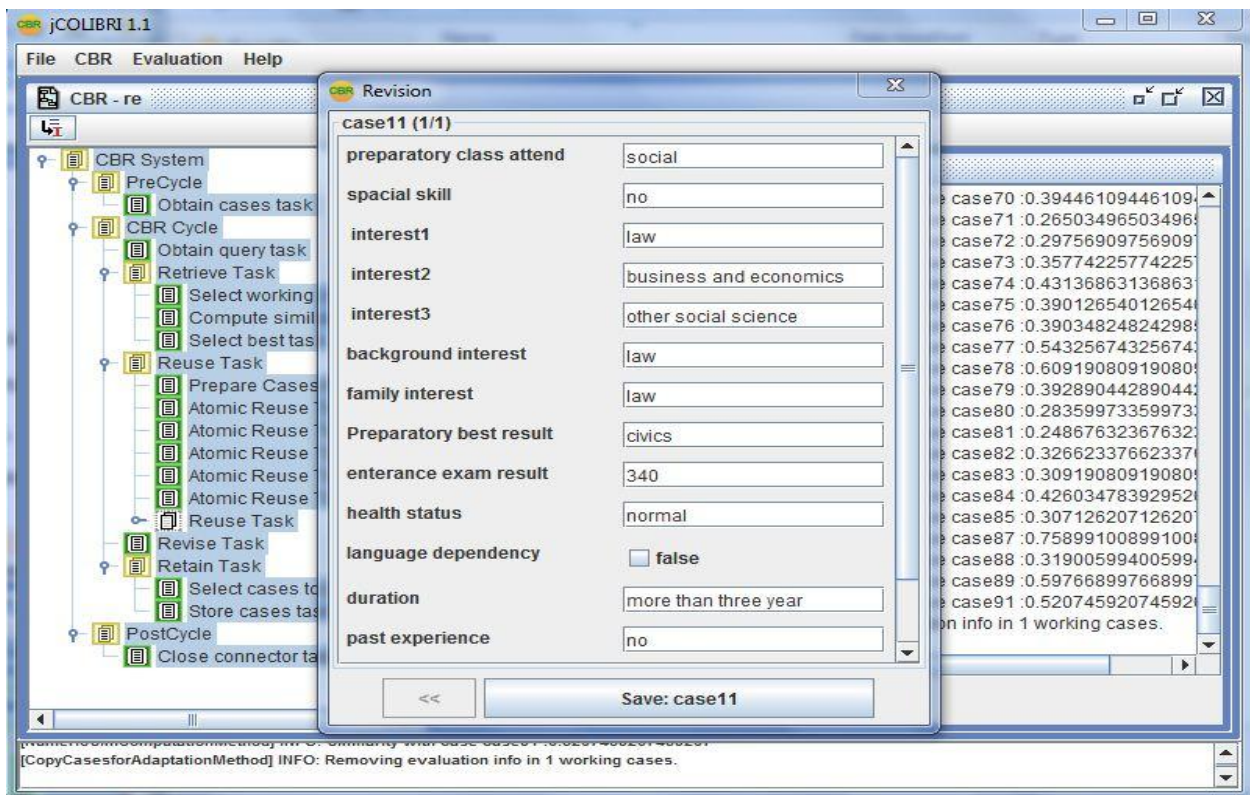


Figure 5. 2 case revision testing

5.2.3. Case Retaining testing

The last cycle in JCOLIBRI framework is retaining which is an important step to store new cases together with the existing case. These new cases are used as an input or case for the next field of study selection process. There might be change of criteria with National Agency for

Examination such as entrance exam result of the year. At this time, new cases which are recommended to the students and used by current students in the field of study selection process can be stored and reused by students after some days later. This can be processed with retain stage after revision case. Manual indexing is important to assign the index value of new case as shown below in figure 5.3.

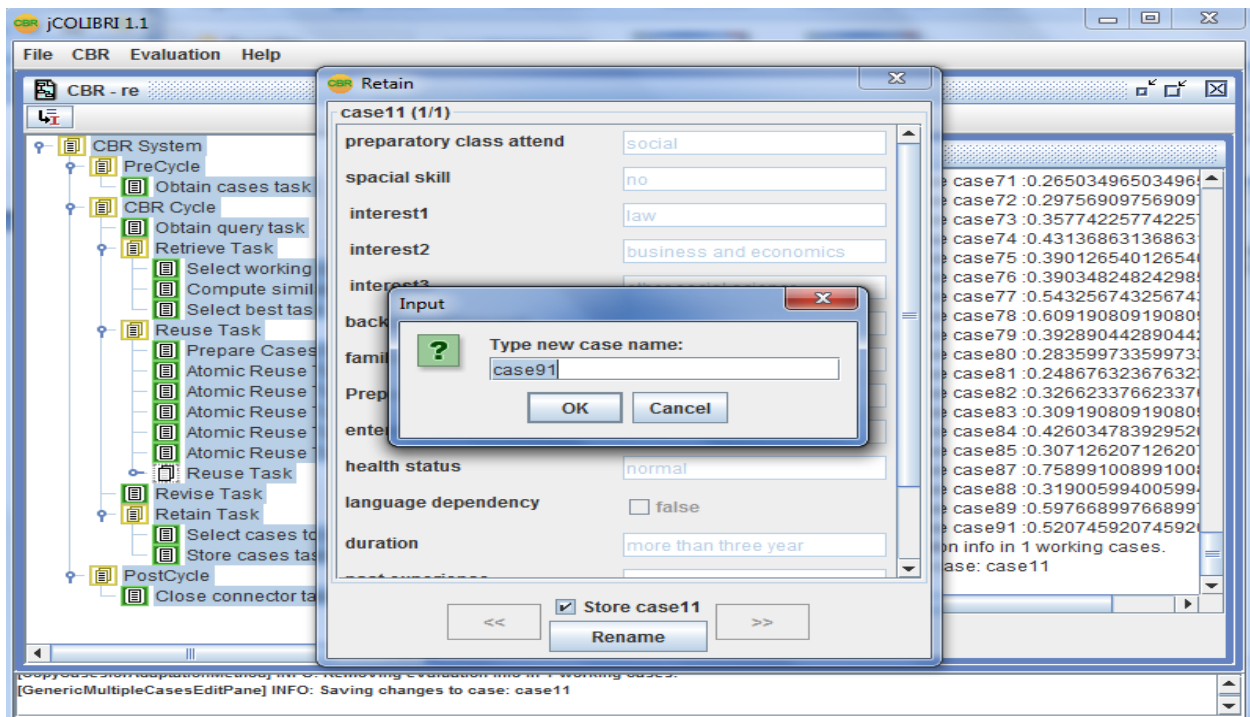


Figure 5. 3 Case Retaining windows

5.2.4. User Acceptance Testing

The validity of the case base recommender system is tested using user feedback to check its applicability in field of study selection process. The potential users of the system are students who are completed their preparatory class and fulfilled the minimum entrance exam requirement of the year. To evaluate the applicability of the prototype, feedback was collected from students who have joined higher education institutes and domain experts. These students were selected randomly from Debre Markos University and experts were selected purposely from each field of study to evaluate the acceptance and performance of the case base recommender system. Since the system need installation of the JCOLIBRI software, only twenty

students and six domain experts from the different field of study were selected to evaluate the system.

Each group tested the system using their own test queries and evaluates the search results in terms of clarity and adequacy to retrieve the existing cases and recommendations based on the query. The criteria for evaluation were as listed in the following table and students and experts were asked for each case to assign values based on scale value and respective marks for each value. These values and respective marks were; poor=1, fair=2, good=3, very good=4 and excellent=5. The following table shows the general criteria for evaluations of the system.

Evaluation criteria	Performance Value					
	1	2	3	4	5	Average
Adequacy and clarity of the system in recommendation process			1	3	2	4.2
Relevancy of the attributes in representing field of study selection process				2	4	4.7
Is the user interface of prototype interactive			1	4	1	4
Ease of use				5	1	4.2
Relevance of retrieved cases in the decision making to support recommendation system				3	3	4.5
Is the system efficient in time and memory		1	3	1	1	3.3
Fitness of the final solution to the new case		2	3	1		2.8
Significance of the system in domain area			2	3	1	3.2
	Total Average					3.86

Table 5. 3 performance evaluation by domain expert

As shown in table 5.3, 16.67 % of the domain expert respondents' rate adequacy and clarity of the system as good, 50% rate as very good and the remaining 33.33% of the respondents rate as excellent.

At the same time, relevancy of attributes to represent field of study process have been rated by domain experts as 33.33% very good and 66.67% as excellent. In the case of interactivity, 16.67% of respondents rate the interactivity of user interface as good, 66.67% rate as very good and 16.67% rate as excellent. In the case of easy use of the system, 83.3% of the

respondent rate the system as very good and 16.67% of them rate as excellent. Similarly, relevance of the retrieved cases in decision making to support users by recommending most similar cases rated as very good by 33.33% of the respondents whereas the remaining 66.67% of the respondents rate it as excellent. 16.67% of the respondent rates the efficiency of the system in time and memory as fair, 50% of the respondent rate as good, 16.67% them as very good and the remaining 16.67% as excellent. In the case of fitness of the final retrieved solution to the new problem at hand, 33.33% of the respondent rate as fair, 50% of them rate as good and 16.67% of the respondent rate the fitness of final retrieved solutions as very good. The last evaluation parameter which deals about the applicability of the system to domain area also evaluated. 33.33% of the respondent rate the applicability of the prototype in their domain area as good, 50% rated as very good and the remaining 16.67% of the respondent rate as excellent.

As shown from the results in the above table 5.3, domain experts assigned less value for the last three criteria. This less rating might be the result of testing query used by domain experts.

Finally, based on the evaluation of all the domain experts, the average performance of the prototype is 3.86 which indicate the performance of the system is approximately very good. From this performance, the researcher deduced that the prototype of the system has promising applicability in field of study selection process.

On the other hand, table 5.4 below shows the performance evaluation of the prototype by the students.

Evaluation criteria	Performance Value					
	1	2	3	4	5	Average
Adequacy and clarity of the system in recommendation process			11	6	3	3.6
Relevancy of the attributes in representing field of study selection process			4	14	2	3.9
Is the user interface of prototype interactive				15	5	4.25
Ease of use			7	12	3	4.2
Relevance of retrieved cases in the decision making to support recommendation system			14	4	2	3.4
Is the system efficient in time and memory				7	13	4.6
Fitness of the final solution to the new case			12	7	2	3.7
Significance of the system in domain area				13	7	4.35
	Total Average					4.01

Table 5. 4 performance evaluation of the system by students

As shown in table 5.4, 55% of the respondents rated adequacy and clarity of the system as good, 30% of them rated as very good and 15% of the respondent rated as excellent. In the case of relevancy of attributes for decision making, 20% of the respondents rated as good, 70% rated as very good and the remaining 10% of the respondents rated as excellent. The system`s user interface also evaluated by the students as very good and excellent with 75% and 25% of the respondents respectively. Similarly, 35% of the respondents rated the ease use of the system as good, 60% of the respondents rated as very good and only 15% of them rated as excellent. Relevance of retrieved cases in the recommendation process for decision making rated as good by 70% of the respondents, very good by 20% of them and the remaining 15% of respondents rated as excellent. Again, 35% of the respondents rated the efficiency of system in time and memory usage as very good and 65% of them rated as excellent. In the case of Fitness of the final solution to the new case, 60% of the respondents rated as good, 35% of them rated as very good and the only 10% of the respondents rated the fitness of system to the new case as excellent.

When the researcher tried to compare the performance value of the system, some criteria have different value by domain expert and students. For instance, adequacy and clarity of the system

in recommendation process rated 3.6 by students and 4.2 by domain experts. Since domain experts have more understanding about the system, they know more about adequacy and clarity of the recommender system in field of study selection process. Relevant attributes were selected by domain experts. Therefore, domain experts know more about these attributes rather than students and they rated higher value. In the case of efficiency of system in time and memory as well as significance of the system, immediate users (students) rated higher value than domain experts. This is because of proximity of the problem in selecting field of study process. Students know about the problem of field of study selection process to decide which field of study fits their interest and background information.

Generally, the significance of the system rated as very good by 65% of the respondents and excellent by 35% of them. The average performance of the system is 3.86 (77.2%) and 4.01 (80.2%) by the domain experts and students respectively. When the researcher compares these two results, the performance of the system scored by students is higher than domain experts. This shows that the system is very good with its performance and more acceptable and applicable for the students. This indicates that the system has achieved its goal to give recommendation to students in their field of study selection process.

CHAPTER SIX

CONCLUSION AND RECOMMENDATION

6.1. Conclusion

Recommender systems have proven to be an important response to the information overload problem by providing users with more proactive and personalized information services that help people to find items of interest from large information spaces. The main goal of this research is to develop a prototype case based recommender system that can assist the students in choosing their field of study to join governmental higher institution in Ethiopia. For this purpose the knowledge was acquired from domain experts and students who are successful with their best choice of field of study with interview.

Both tacit and explicit knowledge for the study was acquired from domain experts and case features were formulated. Relevant attributes in the form of case structure which will have direct impact for decision was also identified. Case base recommender system was employed in representing the necessary knowledge required for handling field of study selection process. The case base was developed using JCOLIBRI case based framework tool which is the most compatible and reliable tool to develop recommender system.

The reason for applying case base recommender system is that rule base system is not efficient in creation of knowledge base when there are different reasons at different time. The main problem of rule base recommender system is changing and updating of knowledge base from time to time. Rule base recommendation will not update itself to reuse the new criteria since it is only based on specific rules and facts which are restricted. The main criteria applied in field of study selection process can be changed from time to time, for instance students` result which is the main determinant factor to place students in their first choice will vary from year to year. Case base recommender system use past experiences as the domain knowledge and can often provide a reasonable solution through appropriate adaption. Since the system stores the new case within the existing cases, new case can be used as a case base for the next time.

The retrieval task starts with a new case description and ends with best matching similarity case of field of study selection process that have been found in the case base. As the students enter the query through user interface window, the system searches the best matching cases from the case base and return possible solution by measuring similarity of existing cases and new query using nearest neighbor algorithm. If exact matching occurs between existing and new queries, students can directly derive the solution; whereas the similarity matching is partial, reuse and revision tasks can be done to make the solution best. Finally, recommendation results that are best solution to the problem are stored to the case base for later use. The system learns incrementally in such a way and stores every new case to the case base. Students should use appropriate query which is similar to the cases in the case base to get relevant recommendation. Using inappropriate query will lead wrong recommendation and students can make wrong decision to select field of study.

There are potential benefits in using field of study selection recommender system for the students to make decision easily when they select field of study. This system considers different criteria to give recommendation for the student that enables them to choose field of study easily based on recommendations given by the system. The system will overcome information overload that comes from different directions such as peer and family about field of study selection and they can choose based on the previously solved cases. Students can choose best field of study based on their interest, special skill, background interests and others identified by experts. Therefore they might be satisfied with their choice and can be successful in their higher education.

The prototype of the system was tested by comparing the solution provided by the system with previously recommended cases by selecting human experts from each field of study from Debre Markos University and students in the university. The result shows most similar cases were retrieved when attribute of the new case was similar to the cases in the case base. All the useful cases which are important for decision making with the students were included in partial matching cases.

Finding the relevant attributes and cases was the challenging task in this study. The researcher tried to collect these attributes from domain experts. Domain experts have been selected main attributes based on their personal view and background information. At the same time, collecting the case also performed based on information of successful students from the higher education institutions manually.

In this study, the performance of the system has been measured by collecting feedbacks from few domain experts and students manually. Testing the system with test cases and user acceptance testing method was the challenging task for this study.

6.2. Recommendation

Although the results of this study are necessary, there are a number of problems to be investigated by future researchers and further work needs to be done in order to upgrade the prototype system to a real system. A continuation of this research also can assume on extending the scope of the system to develop recommender system that assists students by adapting the retrieved precedent cases. This will increase the usability of the system in assisting students when they select field of study.

The following recommendations are made based on the findings and limitations of this research.

- The researcher tried to identify main problems of students when they select field of study to join higher institution. Even though there are problems to select specific departments after joining university, this study tried to overcome the problem student face when they select band and sub-bands. Based on this, the researcher recommended for future researchers to investigate other specific problems such as department selection problem after they joined in to different band or sub-band.
- Selection of main attributes that are important to case base recommender system was also challenging task. This is because the selection process was done manually with the help of domain experts by considering which attributes are more important for decision making in selecting field of study. The researcher suggests applying machine learning system to generate best attributes by ranking.
- The use of nearest algorithm increases linearity of retrieval time when there are many cases and it returns the nearest match even with dissimilarity cases in the source and new case. In the future it is recommended to use other retrieval algorithms such as template retrieval that returns all cases that fits certain parameter.
- The researcher has represented few cases in the knowledge base. This is due to that some departments are not found in the nearby university to the researcher such as theatrical and fine art. The system can be more accurate if all field of studies and specific departments included under the case.

- This recommender system recommends only one recommendation even if other cases which are similar to the new case are displayed in ranked order. For the future work, the researcher recommended to use other tools that can recommend more alternative solutions.
- Applying automatic performance measurement system will increase the performance of the system.
- The developed prototype recommends field of study that is important for the students to make decision easily. The researcher recommends for the future work to improve this system that can recommend university as well as field of study selection by providing detail information about each selection and finally, National Agency for Examination also recommended to consider their choice that is supported by this case base recommender system.
- The performance of the system can be improved if hybrid approach is employed by combining rules, cases and models since these rules, cases and models have complementary strength. For the future, it is better to integrate these approaches to make knowledge base system more successful.

REFERENCES

1. Abraham, A. (2005). Rule based Expert Systems. Oklahoma State University, Stillwater, OK, USA.
2. Adomavicius, G. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, *IEEE Trans. Knowledge and Data Eng.*, vol. 17, no. 6, pp. 734-749.
3. Adomavicius, G. and Tuzhilin, A. (2005). Query Language for Recommender Systems. Stern School of Business, New York University.
4. Amodt, A. and Plaza, E. (1994). Case-based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *AI Communications*.
5. Aradhya, A. (2001). Survey on Recommender Systems.
6. Asanov, D. (2011). Algorithms and Methods in Recommender Systems; Berlin Institute of Technology, progetto didattico. Berlino, Germania.
7. Baeza-Yates, R., and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: CM Press.
8. Balabanovic, M. and Shoham, Y. (1997). Content-based, collaborative recommendation. *Communications of the ACM*, 40(3).
9. Bamshad, M. (2002). Recommender Systems: ECAI'08 Workshop on Recommender System.
10. Barry, S. (2009). Case-Based Recommendation. *The Adaptive Web Methods and Strategies of Web Personalization*, University of Warwick.
11. Belkin, N. and Croft, B. (1992). Information Filtering and Information Retrieval, *Comm. ACM*, vol. 35, no. 12, pp. 29-37.
12. Ben Schafer, J. Dan Frankowski, Herlocker, J. and Sen, S. (2007). Collaborative Filtering Recommender Systems. Department of Computer Science, University of Northern Iowa.
13. Ben Schafer, Joseph Konstan, J. Riedl, J. (2001). Recommender Systems in E-Commerce, GroupLens Research Project, University of Minnesota.

14. Bendakir, N. and Imeur, E. (2006). Using Association Rules for Course Recommendation. American Association for Artificial Intelligence.
15. Burke, R. (2002). Hybrid recommender systems: survey and experiments. *User Modeling and User-Adapted Interaction* 12, no. 4.
16. Burke, R. (2006). Knowledge based recommender systems, University of California, Irvine.
17. Burke, R. (2007). Hybrid Web Recommender Systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.). *The Adaptive Web: Methods and Strategies of Web Personalization*.
18. Cai-Nicolas Ziegler, Sean McNee, M. Konstan, A. Lausen. G (2005). Improving Recommendation Lists Through Topic Diversification; GroupLens Research, Univ. of Minnesota. China.
19. Candillier, LMeyer, . F. Fessant, F. and Jack, K. (2009). State-of-the-art recommender systems.
20. Cantador, A. Bellogin, D. Valet, (2010). Content-based Recommendation in Social Tagging Systems. Department of Engineering, Universidad Autónoma de Madrid.
21. Chau, M. (2002). Design and Evaluation of a Multi-agent Collaborative Web Mining System. In *Decision Support Systems*.
22. Chen, Y. L., Hsu, C. L., & Chou, S. C. (2003). Constructing a multi-valued and multi-labeled decision tree. *Expert Systems with Applications*, 25, 199–209.
23. Clancey, W. (1984). Knowledge Acquisition for Classification Expert Systems. *Proceedings of ACM National Conference on computer science*, Stanford University.
24. Clarke, S., & Willett, P. (1997). Estimating the recall performance of search engines.
25. Ebrahim Kassa, (1999). An Expert Based Admission and Placement System. School of Information Science, Addis Ababa University, Addis Ababa, Ethiopia.
26. E. Sajja & Akerkar, (2010). Knowledge Based Systems for Development. *Advanced Knowledge Based Systems: Model, Applications & Research*. Vol. 1, pp 2 – 11.
27. Ekdahl, M. Lindstr and Svensson, C. (2002). A student course recommender: Master of Science Programme, Lulea University of Technology, Department of Computer Science and Electrical Engineering / Division of Computer Science and Networking

28. Fabiana, L. and Francesco, R., (2003). Case-based recommender systems: A unifying view, in Intelligent Techniques for Web Personalization, IJCAI 2003 Workshop, p. 2-10.
29. Fong, S. and Biuk-Aghai, R. (2009). An Automated Admission Recommender system for Secondary School Student. The 6th international conference on Information Technology and application.
30. Getachew Feleke, (2008). Higher education entrance student placement processing and retrieval system: project paper: Addis Ababa University School of Computer Science.
31. Giovanni Semeraro, (2010). Content based Recommender Systems. problems, challenges and research directions, Department of Computer Science University of Bari.
32. Goldberg, Eigentaste, (2000). A Constant Time Collaborative Filtering Algorithm. Available at: www.ieor.berkeley.edu/goldberg/pubs/eigentaste.pdf.
33. Grzegorz, J. Nalepa, (2007). Prototype Prolog API for Mindstorms NXT. Institute of Automatics, AGH University of Science and Technology, Poland.
34. Guy S., Ronen I., Brafman and David Heckerman (2005). An MDP-Based Recommender system: Recommendation as a Stochastic Sequential Decision Problem. Journal of Machine Learning Research; pp 1267-1269.
35. Guy, S. and Asela, G. (2010). Evaluating Recommendation Systems: Introduction to Recommender System Handbook, pp. 257-265.
36. Haruechaiyasak, C M.-L. Shyu, and S.-C. Chen, (2004). A Data Mining Framework for Building A Web-Page Recommender System. IEEE International Conference on Information Reuse and Integration (IRI'2004), Las Vegas, Nevada, USA, pp. 357-262.
37. Herlocker, J. (2000). Understanding and Improving Automated collaborative Filtering Systems. University of Minnesota: Faculty of Graduate School. Available at: <http://cs.oregonstate.edu/~herlock/papers.html>
38. Imran, G. and Aniza, A. (2010). Measuring learner's performance in e-learning recommender systems. Australasian Journal of Educational Technology, University of Malaya.
39. Ivan Cantador Gutiérrez, Madrid, (2008). Exploiting the conceptual space in hybrid recommender systems. A semantic-based approach, Dissertation, Madrid.

40. Janet L. Kolodner, (1992). An Introduction to Case-Based Reasoning. Artificial Intelligence Review 6. College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0280, U.S.A.
41. Jannach, D. Zanker M. and Friedrich G., (2011). Recommender Systems: an introduction. International Joint Conference on Artificial Intelligence. Alpen-Adria Universität Klagenfurt.
42. Linden, G. Smith, B. and York, J. (2003). Amazon.com recommendations. "Item-to-item collaborative filtering," IEEE Internet Computing, vol. 7, pp. 76–80. Available: <http://portal.acm.org/citation.cfm?id=642462.642471>.
43. Luan, J. (2002). Data Mining Application in Higher Education: SPSS Executive Report.
44. Makfi, P. (2011). Introduction to knowledge modeling and neural network. Available at: http://www.makafi.com/KCM_intro.html.
45. Markus Zanker, (2010). Introduction to Recommender Systems, Tutorial at ACM Symposium on Applied Computing, Switzerland.
46. Michael, J. Pazzani and Billsus, D. (2007). Content-Based Recommendation Systems. Springer-Verlag Berlin Heidelberg, Rutgers University.
47. Min-Ling Zhang and Zhi-Hua Zhou (2004). A k-Nearest Neighbor Based Algorithm for Multi-label Classification. National Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China.
48. Ministry of Education (2011). Manual of higher education institution student's placement process.
49. Nikos, Manouselis, G. and Constantina, Costopoulou, I. (2006). Designing a Web-based Testing Tool for Multi-Criteria Recommender Systems. Advance online publication, University of Athens.
50. Ning Xu, (2007). Explanation Interfaces in Recommender Systems. Information and Communication Technology in Business. Leiden Institute of Advanced Computer Science, Leiden University.
51. O'Connor, M., Cosley, D., Konstan, J.A., Riedl, J. PolyLens, (2001). A Recommender System for Groups of Users. In Proceedings of ECSCW Bonn, Germany p. 199-218

52. P. Symeonidis, P. Nanopoulos, A. Apostolos, N. Papadopoulos, Manolopoulos, Y. (2007). Collaborative recommender systems: Combining effectiveness and efficiency.
53. Pazzani, M. J., & Billsus, D., (2007). Content-Based Recommendation Systems. The Adaptive Web. Lecture Notes in Computer Science vol. 4321, 325-341.
54. Pennoch, David, M. and Eric Horvitz, (2000). Collaborative Filtering by Personality Diagnosis: A hybrid memory- and model-based approach. In Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence. pages 473-480. Available at: <http://citeseer.nj.nec.com/pennock00collaborative.html>.
55. RC Chakraborty, (2010). Artificial intelligence; Knowledge Representation, Issues, predict logic, rules. Available at: http://myreaders.info/html/artificial_intelligence.html.
56. Russel, S. J. and Norvig, P. (2010). Artificial Intelligence. A Modern Approach. 3rd ed. Prentice Hall.
57. Satyanarayana, K. and Rajagoplan, S.P. (2007). Recommender system for Educational Institutions. Asian Journal of Information Technology. M.G.R University, India.
58. Shadrack, S.B., Iussier, J.W and Hinkle, R., (2005). Concept Development for Future Domain: A new Method of Knowledge Elicitation. U.S Army Research Institute for the Behavioral and Social Science, Technical reports No 1167.
59. Shardanand, U., and Maes, P. (1995). Social Information Filtering: Algorithms for Automating 'Word of Mouth'. In Proceedings of CHI '95. Denver, CO. Available at: <http://citeseer.nj.nec.com/shardanand95social.html>
60. Shimazu, H. (2002). Expertclerk: A Conversational Case-based Reasoning Tool for Developing Salesclerk Agents in E-Commerce Webshops. Artificial Intelligence Review 18 (3-4), 223–244.
61. Simon, F., Robert, P. (2009). An Automated University Admission Recommender System for Secondary School Students: The 6th International Conference on Information Technology and Applications.
62. Sun Z. and Luo, N. (2010). A new user-based collaborative filtering algorithm combining data distribution, in Proceedings of the 2010 International Conference of Information

Science and Management Engineering. Washington, DC, USA. Volume 02, pp. <http://dx.doi.org/10.1109/ISME.2010.48>.

63. Terveen, L. and Hill, W. (2001). Beyond Recommender Systems. Helping People Help Each Other.
64. Toffler, A. (1981). The Third Wave. London, Pan Books Ltd.
65. Ulle Endriss, (2000). An Introduction to Prolog Programming: Lecture Note. King's college, London. V.1. pp. 2-21.
66. Vialardi, J. Bravo, L. Shafti, A. Ortigosa, (2009). Recommendation in Higher Education Using data Mining Techniques. Educational Data mining.
67. Vozalis, E. Margaritis, G. (2001). Analysis of Recommender Systems' Algorithms. Applied Informatics, University of Macedonia, Thessaloniki, Greece.
68. Watson, I. (1999). Case-Based Reasoning is a Methodology not a Technology. *Journal of Knowledge-Based Systems*.

Appendix A: Attribute and case collection

A. Interview questions to collect main attributes

The main objective of this questioner is to collect basic information about field of study selection process. The researcher will develop **case base recommender system** that advices students in their field of study selection process based on the collected cases.

1. faculty-----
2. Which criteria are most important to assign students in your department?(you can choose all or some of them)
 - a. Age
 - b. Sex
 - c. Interest 1,2,3
 - d. Background interest
 - e. Family interest
 - f. Result
 - g. Relate subject result in preparatory class
 - h. Preparatory class attend
 - i. Entrance exam result
 - j. Health condition
 - k. language
 - l. duration
 - m. past experience /worked on related field
 - n. need of nearby university to family
 - o. economic back ground
 - p. employment
 - q. pre request subject taken at preparatory level

Additional criteria

3. Is there need of Language dependency
 - a. Yes
 - b. no
4. Special skill required for this department

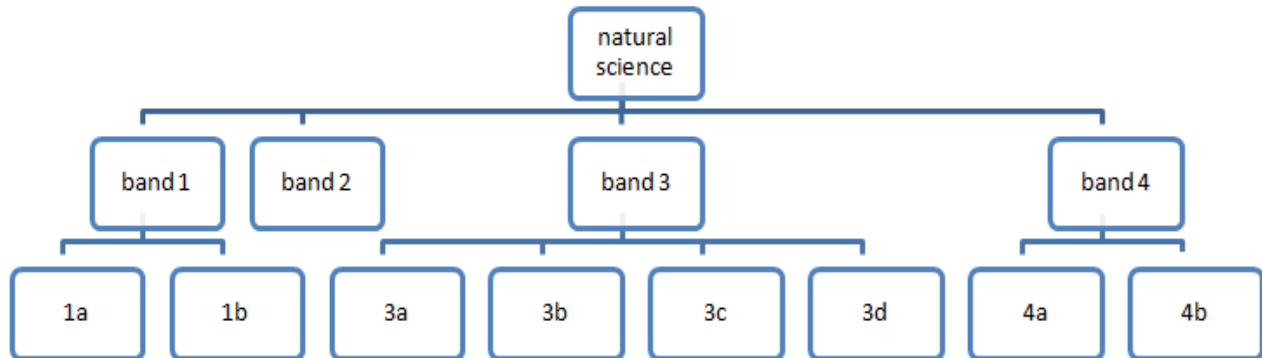
Appendix B: sample cases

cases	preparator	spacial skill	interest1	interest2	interest3	background	family inter	Preparator	enterance	health stat	language re	duration	past experi	recommen	band
case1	natural	drawing	Engineering	computer s	health	Engineering	Engineering	physics	312	normal	no	more than	yes	Engineering	1a
case2	natural	computer	Engineering	computer s	Natural an	computer s	computer s	IT	453	normal	no	more than	no	computer s	1b
case3	natural	computer	computer s	Engineering	Natural an	computer s	Engineering	IT	366	normal	no	more than	no	computer s	1b
case4	natural	drawing	health	Engineering	computer s	Engineering	medicine	maths	356	normal	no	more than	yes	Engineering	1a
case5	natural	drawing	Engineering	veterinary	computer s	Engineering	Engineering	physics	505	normal	no	more than	yes	Engineering	1a
case6	natural	drawing	Engineering	health	computer s	Engineering	medicine	physics	512	normal	no	more than	yes	Engineering	1a
case7	natural	computer	medicine	veterinary	computer s	medicine	Engineering	biology	456	normal	no	more than	yes	medicine	3a
case8	natural	computer	computer s	engineering	medicine	computer s	medicine	IT	450	normal	no	more than	no	computer s	1b
case9	natural	no	medicine	computer s	enegeerini	medicine	medicine	drawing	530	normal	no	more than	no	computer s	1b
case10	natural	computer	computer s	enegeerini	medicine	computer s	computer s	maths	468	normal	no	3 year	no	computer s	1b
case11	natural	drawing	computer s	enegeerini	medicine	Engineering	engineering	maths	410	normal	no	more than	no	Engineering	1a
case12	natural	no	Engineering	medicine	computer s	enegeerini	medicine	physics	360	normal	no	more than	no	Engineering	1a
case13	natural	no	Natural an	Engineering	medicine	chemistry	biology	chemistry	510	normal	no	3 year	yes	Natural an	2
case14	natural	no	Natural an	Engineering	medicine	health	biology	chemistry	290	normal	no	3 year	yes	Natural an	2
case15	natural	no	Natural an	Engineering	computer s	medicine	medicine	biology	290	normal	no	3 year	yes	Natural an	2
case16	natural	no	Engineering	Natural an	computer s	Natural an	medicine	chemistry	345	normal	no	3 year	yes	Natural an	2
case17	natural	no	Natural an	computer s	Engineering	computer s	medicine	maths	298	normal	no	3 year	yes	Natural an	2
case18	natural	no	computer s	Natural an	medicine	Natural an	medicine	physics	411	normal	no	3 year	yes	Natural an	2
case19	natural	no	natural an	other agric	enegeerini	natural an	computati	enegeerini	390	normal	no	more than	no	natural an	2
case20	natural	no	medicine	public heal	other heal	medicine	medicine	biology	540	normal	no	more than	yes	medicine	3a
case21	natural	no	public heal	medicine	dental scie	medicine	medicine	biology	468	normal	no	more than	no	public heal	3b
case22	natural	no	other heal	public heal	medicine	other heal	medicine	chemistry	490	normal	no	more than	no	other heal	3c
case23	natural	no	dental scie	veterinary	public heal	dental scie	public heal	maths	378	normal	no	more than	no	dental scie	3d
case24	natural	no	public heal	medicine	other heal	public heal	medicine	chemistry	502	normal	no	more than	yes	public heal	3b

case80	social	no	other socia	business a	law	business a	law	geography	389	blind	yes	3 year	yes	other socia	6b
case81	social	music	Theatrical,	other socia	law	music	music	sport	423	blind	no	3 year	no	Theatrical,	6c
case82	social	no	law	Business a	other socia	law	law	civics	423	normal	no	more than	no	law	6a
case83	social	no	other socia	business a	law	business a	law	geography	390	blind	yes	3 year	yes	law	6a
case84	social	sport	Theatrical,	other socia	business a	music	music	biology	289	blind	no	3 year	no	Theatrical,	6c
case85	social	no	Business a	law	other socia	Business a	law	economics	292	normal	no	3 year	no	Business a	5
case86	social	music	other socia	Theatrical,	law	music	business a	geography	270	blind	no	3 year	no	Theatrical,	6c
case87	social	no	business a	law	other socia	business a	business a	economics	456	normal	no	3 year	yes	Business a	5
case88	social	no	business a	other socia	law	business a	law	civics	282	normal	no	3 year	no	business a	5
case89	social	no	other socia	business a	law	other socia	law	geography	243	blind	yes	3 year	yes	other socia	6b
case90	social	music	other socia	Theatrical,	law	music	law	sport	288	blind	no	3 year	no	Theatrical,	6c
case91	social	music	Theatrical,	Business a	other heal	art	music	biology	343	blind	no	3 year	no	Theatrical,	6c
case92	social	no	other socia	business a	law	business a	law	geography	289	blind	yes	3 year	yes	other socia	6b
case93	social	music	law	Theatrical,	business a	Theatrical,	law	biology	451	blind	no	3 year	no	Theatrical,	6c
case94	social	no	other socia	Business a	law	business a	law	civics	320	normal	no	3 year	no	business a	5
case95	social	no	business a	law	other socia	other socia	law	geography	299	blind	yes	3 year	yes	business a	5
case96	social	art	other socia	Theatrical,	law	art	business a	sport	310	blind	no	3 year	no	Theatrical,	6c
case97	social	sprot	Theatrical,	Business a	other heal	Theatrical,	art	biology	302	blind	no	3 year	no	Theatrical,	6c
case98	social	no	law	business a	other socia	law	business a	geography	340	normal	FALSE	3 year	no	law	6a
case99	social	no	law	business a	other socia	law	law	civics	340	normal	FALSE	more than	no	law	6a
case100	social	no	law	other socia	business a	law	business a	general bu	342	normal	no	3 year	no	other socia	6b
case101	social	no	other socia	business a	law	business a	law	geography	530	physically h	no	3 year	no	other socia	6b
case102	social	no	law	business a	other socia	law	business a	history	412	normal	no	more than	no	law	6a
case103	social	no	business a	law	other socia	background	law	civics	361	normal	no	3 year	no	business a	5
case104	social	music	Theatrical,	law	other socia	Theatrical,	law	history	341	normal	no	3 year	yes	Theatrical,	6c
case105	social	no	business a	law	other socia	law	law	geography	298	normal	no	3 year	no	law	6a

Appendix C: Structure of field of study at preparatory school with their bands

A. Natural and Computational Sciences



Band1=Engineering and Technology

1a= Engineering science

1b= computer science

Band2= Natural and computational science

Band3= Medicine and Health science

3a= medicine

3b= public health officer

3c= other health science

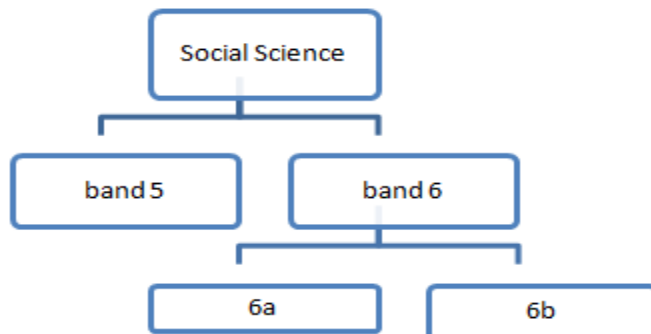
3d= dental science

Band4= Agriculture and Natural science

4a= veterinary medicine

4b= other agriculture and natural resources.

B. Social Science



Band 5= Business and Economics

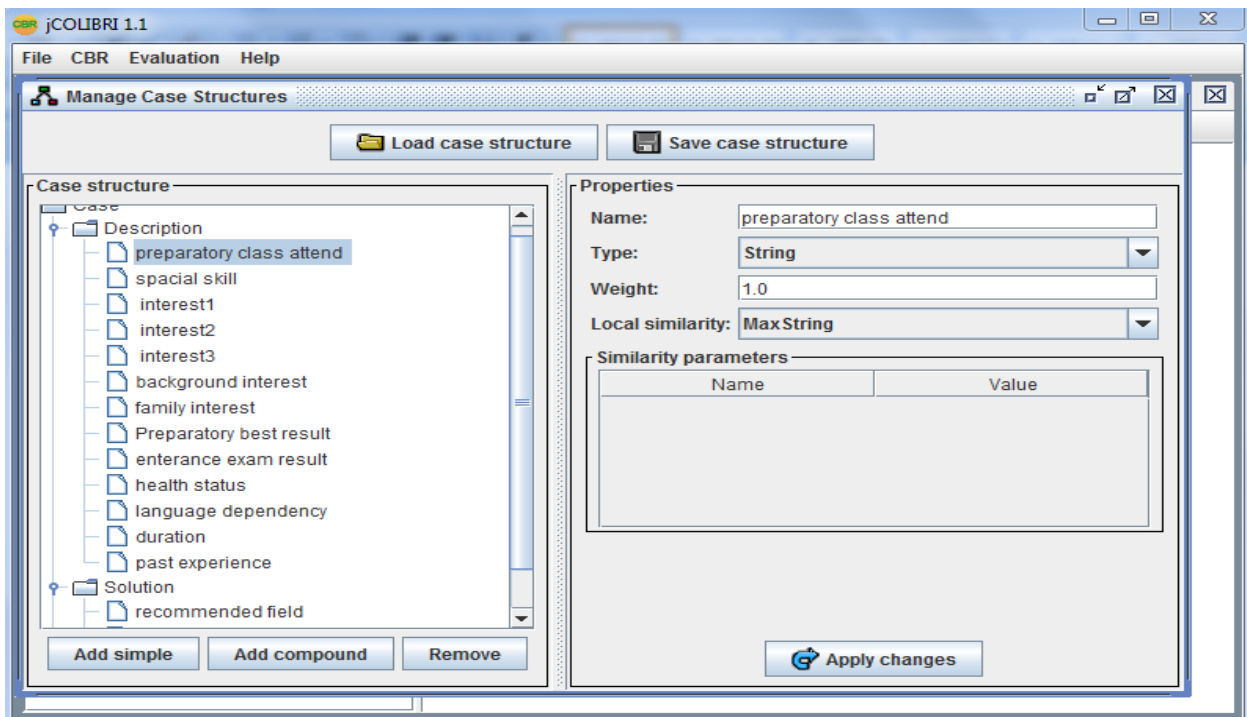
Band 6= Social Science and Humanities

6a = law

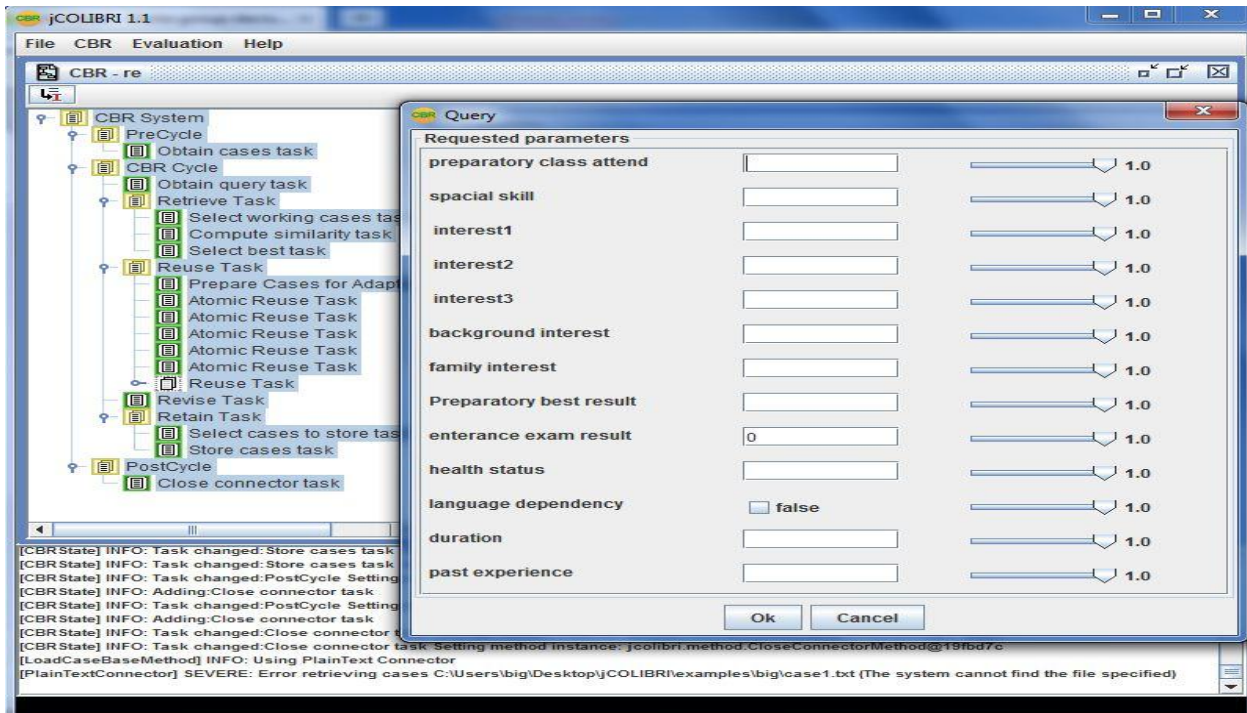
6b= other social science and humanities

Appendix D: JCOLIBRI user Interface

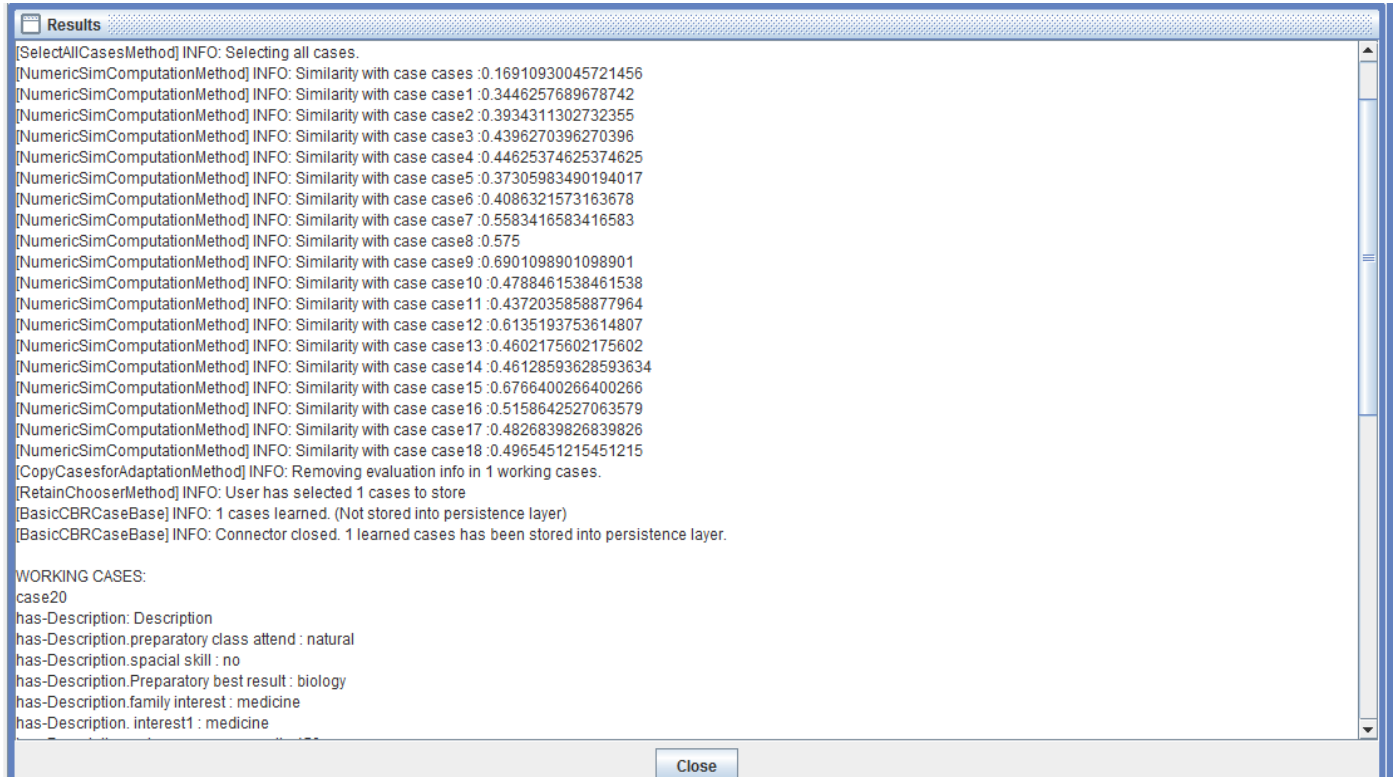
A. Descriptions and solution attributes of recommender system



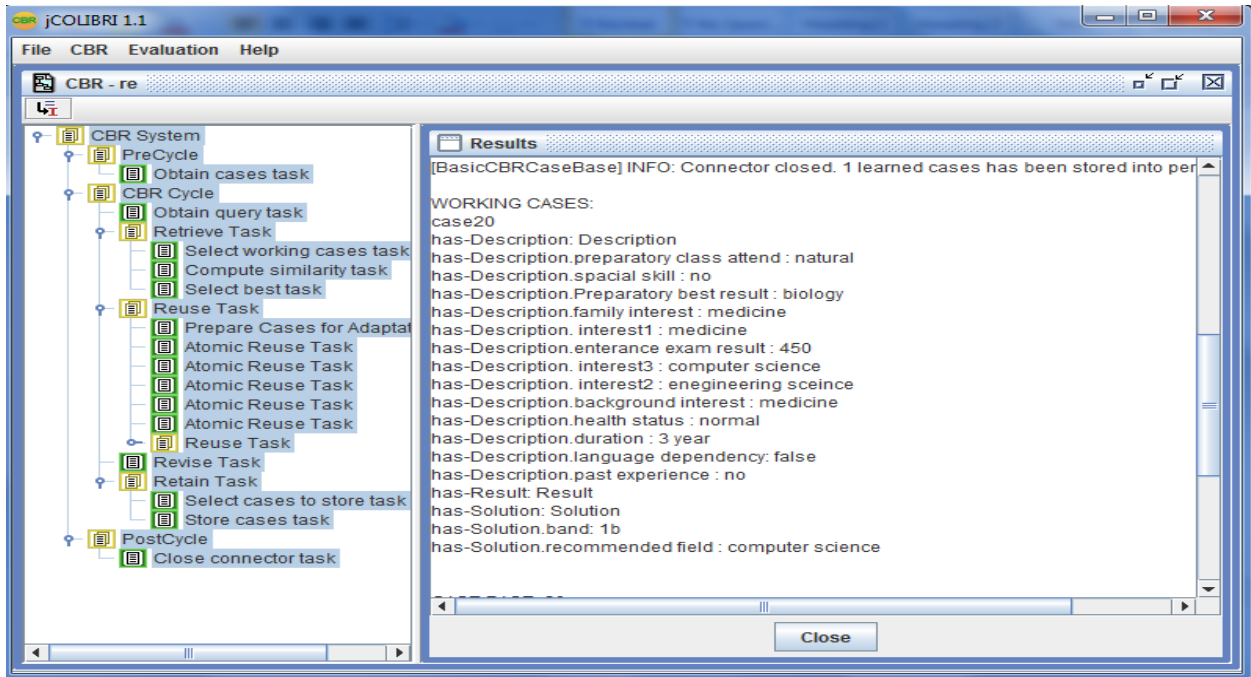
B. User interface used for query entry



C. Sample of case similarity measurement with the new case



D. Sample output after similarity measurement



Appendix E: Evaluation criteria for performance evaluation of the system

A. By domain expert

Evaluation criteria	Performance Value					
	1	2	3	4	5	Average
Adequacy and clarity of the system in recommendation process						
Relevancy of the attributes in representing field of study selection process						
Is the user interface of prototype interactive						
Ease of use						
Relevance of retrieved cases in the decision making to support recommendation system						
Is the system efficient in time and memory						
Fitness of the final solution to the new case						
Significance of the system in domain area						
	Total Average					

B. By students

Evaluation criteria	Performance Value					
	1	2	3	4	5	Average
Adequacy and clarity of the system in recommendation process						
Relevancy of the attributes in representing field of study selection process						
Is the user interface of prototype interactive						
Ease of use						
Relevance of retrieved cases in the decision making to support recommendation system						
Is the system efficient in time and memory						
Fitness of the final solution to the new case						
Significance of the system in domain area						
	Total Average					

Appendix F: list of departments in Debre Markos University used to collect cases

1. Animal science
2. Natural resource management
3. Plant science
4. Rural development
5. Land resource management
6. Accounting
7. Economics
8. Management
9. Banking and finance
10. Nursing
11. Public health
12. Law
13. Biology
14. Chemistry
15. Physics
16. Statistics
17. Mathematics
18. Sport science
19. Ethiopian languages and literature
20. Geography and environmental studies
21. History
22. Psychology
23. Civics
24. Sociology
25. Education
26. English
27. Construction technology
28. Civil engineering
29. Mechanical engineering
30. Electrical engineering
31. Water engineering
32. Information technology