An Ontology Based Personalized Postgraduate Degree Recommendation System

Jayani Jagoda[1], Thushari Silva[2]

University of Moratuwa, Sri Lanka
[1]jayanisamarasekara7@gmail.com, [2]thusharip@uom.lk

Abstract. The quantity and complexity of available information about a specific domain is rapidly increasing on the World Wide Web. Information overloading problem has posed significant challenges on postgraduate studies recommendation. A fresh graduate needs to choose the best postgraduate program that suits her while balancing all her/his requirements including cost, preferred area of study, demanding areas related to their bachelor degree, etc. Current approaches in postgraduate opportunity recommendation systems mine data using traditional data mining techniques hence overlook the semantical data mining. In order to overcome these deficiencies, this research paper presents an ontology based hybrid approach to recommend the most personalized postgraduate degree program which balances matched user's preferences. The proposed approach analyzes the user preferences and user behavior, extracted from the LinkedIn user profiles. An ontology model is also used to generate semantically rich knowledge base to facilitate intelligent recommendation.

Keywords: Recommendation system, Ontology, Content based filtering, Collaborative filtering.

1. Introduction

There is large volume of information about different domains is available on World Wide Web with the rapid development of the technology [1]. Especially the recently emerged social media technologies lead to the era of big data and made the analysis of data from multiple data sources is crucial for highly effective useful recommendations. Recommender systems related to education domain is an example which has been affected from the information overloading system especially generated through social media data. It is a real challenge and time consuming process to filter the necessary information about the postgraduate degree programs from available university web sites. Generally related information has been scattered in different sources and relevant data is flooding through mobile devices at high speed. Integrations of such data sources and mining information to reveal the real insight is vital for decision making. In manual way, a student needs to search the most appropriate degree program by matching their preferences and current qualification level while accessing each platform and reading the syllabus carefully of different postgraduate degree programs. Moreover, others feedback on the quality of the programs which are not readily available on institutional websites play an important role in deciding the postgraduate programs. Therefore, it is essential to harness the real insight of data collected/generated at multiple sources intelligently.

Recommended systems act as a solution for the information overloading problem as it helps to find the mostly related item using information filtering approaches [8]. There are online web applications which suggest postgraduate degree programs to the users whereas the recommendations are not personalized according to the user. Utilization of an ontology in recommendation systems increases the accuracy of recommendations and user satisfaction as ontologies could mine semantics of the keywords hence could improve the matching score. Complete profiling of users are required for highly personalized recommendations [1]. Hence the proposed hybrid approach for postgraduate studies recommendation track the user preference and user behavior in order to improve the accuracy of the recommendations. The user preference can be gained by implicitly or explicitly [6]. In this paper, user preferences and user behavior is recorded using the user's LinkedIn profile data.

There are different approaches taken to address this problem of information overloading in education domain. Most of the solutions contain ontology based recommendation systems whereas the recommendation is not personalized for a specific person. Protégé tool is used for creating ontologies. The user preferences are not semantically analyzed in some of the approaches. There are recommendation systems based on education domain which have used collaborative filtering technique or content based filtering technique. Each filtering technique includes drawbacks. Therefore, the PhD recommendation system uses hybrid filtering technique.

In this paper, a hybrid recommender method based on an ontology is developed. Ontology is used to represent the knowledge for a specific domain [2]. To unify semantically identical but lexically distinct keywords for increasing the matching score an ontology is used in this system. The postgraduate degree courses ontology represents the knowledge of student profile and postgraduate degree profile and their relationships. Protégé tool is used to construct the postgraduate degree courses ontology and number of steps is followed to develop the ontology. First of all, the domain and scope of the ontology is determined. According to the postgraduate degree courses ontology, the domain is postgraduate studies and the scope is narrowed down to master's courses and PhD courses in computer science. Though this study is limited to computer science master's degree programs and PhD degree programs this can be applied to other streams as well with fewer changes. Many ontologies were identified that model higher education domain whereas there were no ontologies to reuse for the intended purpose [3]. Some concepts are further classified in to sub classes which

help to classify the instances of the classes more accurately. The relationships between classes are defined using object properties and data type properties. This recommendation system recommends the most appropriate degree program and a supervisor of a university which is located at user preferred location by referring the user's social network who has done that same degree which is recommended by the system as the final output. Amalgamation of content-based approaches with collaborative filtering methods overcome problems associated with each other. Content-based approaches are good at matching a given textual description with other descriptions in textual form but overlook the wealth of social connections as well other meta data that potentially yield more related results. Hence, this research interconnects collaborative approach as well to capture user connections in the context of recommendation.

The rest of this paper is organized as follows: section 2 presents the related work about the current postgraduate opportunity recommendation systems followed by the proposed approach for the ontology-based personalized recommendation system. The implementation of the system has been presented in Section 4. Section 5 analyzes the results and evaluations while section 6 covers the conclusion of this paper. The rest of this paper is organized as follows: section 2 discusses the related works about the ontology-based personalized recommendation systems. Section 3 explains the proposed approach for the ontology-based personalized recommendation system and the implementation of the system. Section 4 discussed the result analysis of the proposed approach while section 5 discusses the conclusion of this paper. Section 6 is the list of references that are used in this paper.

2. Related work

The process of obtaining information about a specific domain is difficult due to the exponential increase in data generation. Therefore, several approaches are taken by different aspects to provide a suitable solution for this problem. Ontology-based personalized learning path recommendation for course learning [10] research done by Boya Chen just consider learners' scores rated by themselves as the performance scores. Therefore, limitation lies in the performance measurement and more details required to be added in this recommendation system.

The paper [12] suggests a framework for model User Profiles in Personalized Job Recommendation system. This paper has analyzed the domain of job recommendation to identify the relevant concepts and identified the attributes of each concept and the relations between the classes and then implemented an ontology that models user profiles. This ontology can be used for any job recommendation system which is modeled using protégé. This paper has used ontology to model the user profile and provide personalized list of jobs to the user.

The paper [9] which used Social Interaction between friends for knowledge-based Personalized Recommendation system, the research has presented the role of ontologies as sources for the development of recommendation systems. Furthermore, there is an application for the user's social network evaluation between Facebook friends. This research focused on a method that personalized social network-based recommender system with user interest ontology. The user interest ontology is developed for tourism domain with a semantic social recommender system to deal with the problem of less semantic information in personalized recommendation system. In addition, the presented solution can be further developed to a mobile application of the recommendation system which can easily run in a smart phone.

An ontology-based architecture for context Recommendation System in E-learning and Mobile-learning applications, this research [7] is about recommendation system for education domain which provides a context recommendation using ontology in education domain. This paper provides details about recommending the most appropriate learning materials, especially in e-learning and m-learning according to the user preference. This recommendation system use rule based architecture and OWL rules as a recommendation technique.

Neethukrishnan et al [6] proposed a research paper recommendation system predicts the user's item preferences. Ontology based framework is used to overcome the cold start problem and pair wise similarity problem in a traditional recommendation system. Research paper recommendation system suggests the researchers to select the most appropriate research papers. This system has created a personal ontology with the help of reference ontology which has created using protégé. Utilizing the SVM classifier, the semantic similarity between personal ontologies is calculated. This research paper recommendation system recommends the final top ten articles which match the user preferences based on the similarity.

3. Proposed Design and Implementation

3.1 User Profiling by Behavior Analysis Module

This module is responsible for gathering user profile information either explicitly or implicitly. User information includes structured data such as demographic data, academic information and job interests by the user's LinkedIn profile. This information was extracted based on LinkedIn API. The user interacts with the recommendation system through a mobile application. The user has to provide the user's LinkedIn URL for the Login. The user profile module tracks the user behavior such as like for a post, comment on a post, share a post in his or her profile. Text mining techniques are used to extract the relevant user preferences details about education domain by LinkedIn user profile. The user preference discovery can be classified in to three steps.

1. Data extraction

Extract static (structured) data and dynamic (LinkedIn activities) data of the user.

2. Data Integration
Semantic analysis of user activities and integrate the preprocessed features of the user.

3. Interest Discovery
Use data mining techniques to identify the user preferences.

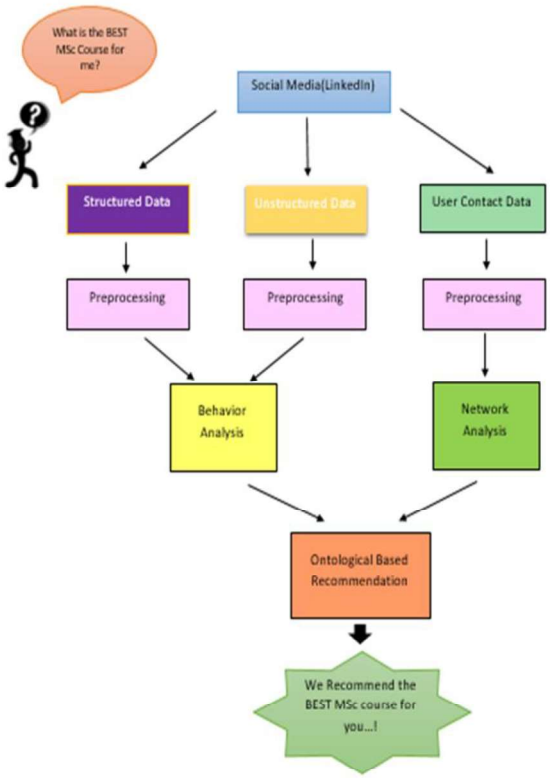

Figure 1: Proposed Design for the recommendation system

3.2 Preprocessing and User Network Analysis Module

This module use LinkedIn contact data of the user to perform social network analysis. Friend network who is currently doing and who have completed postgraduate graduate courses in different countries are identified. People like to follow postgraduate graduate courses if his/her known person has done it previously or if his/her relative lives in that country. The output of network analysis is given as an input to recommendation module. Python Networkx library is used for analyzing network of the user

3.3 Ontology Design for Recommendation Module

- Ontology Design for Postgraduate degree programs

Protégé is used for implementation because it provides facility for the construction of domain ontologies, and there are customized data entry forms to enter data. Protégé allows the definition of classes, class hierarchies, variables, variable-value restrictions, and the relationships between classes and properties of these relationships [4]. In addition, according to some conducted surveys about most frequently used ontology editors and most widely used domain for development of ontology. To define the ontology OWL (Web ontology language) is used. The inputs for the ontology are the user preferences and user's network details.

A. Define the classes
Ontology hierarchy was developed by creating the sub classes after defining the top level classes. The rules and restrictions, object properties and data type properties are modeled which are needed to be embedded when defining the classes and sub classes in the ontology. In the top level class postgraduate degree programs, it is only considered about computer science master degrees and PhD degrees.

There are restrictions for studying postgraduate degree program in a specific field. For example, to obtain a master's degree in computer science student must have followed an Information Technology degree or Computer Science degree as prerequisites. The specialization is recommended according to the analysis of the student's preferences in user profiling. The computer science specialization fields are categorized using the available university web sites and their curricula for postgraduate studies.

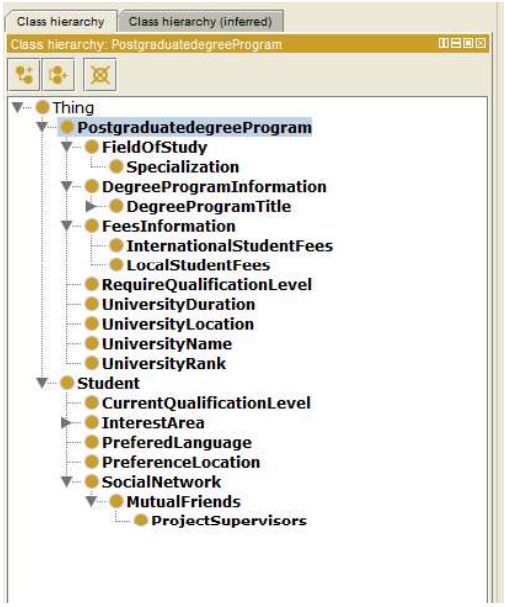

Figure2: Class hierarchy of the Postgraduate degree ontology

B. Define properties of classes
Data and object property restrictions are added to construct the relationship between classes. Table 1 shows the data properties used in the Postgraduate degree program ontology. Table 2 shows the object properties defined in the postgraduate degree program ontology. Individuals are linked by properties from the domain to individuals from the range. Data properties are defined for visualizing the relationship among individuals and literal values.

Table1. Data Properties of the Ontology

Property	Domain	Range
hasDuration_years	UniversityDuration	Integer
hasUniversityLocation	UniversityLocation	String
has Amount	FeesInformation	String
hasUniversityName	UniversityName	String
hasRank	UniversityRank	Integer

Table2. Object Properties of the ontology

Property	Domain	Range
hasMentor	MutualFriends	ProjectSupervisors
isMentorOf	ProjectSupervisors	MutualFriends
hasStudied	MutualFriends	CurrentQualificationLevel
hasFriend	Student	SocialNetwork
Knows	Student	MutualFriends
Recommends	InterestArea	DegreeProgramTitle
hasQualification Level	Student	CurrentQualificationLevel
specializedIn	DegreeProgramTitle	Specialization
hasInterest	Student	InterestArea

C. Define Instances

A set of individuals are defined under each class in order to check the accuracy of the relationships and confirm the ability of the ontology to provide correct answers for queries.

Figure3: Instances of the Postgraduate degree ontology

3.3.1 Recommendation Engine for the Recommendation Module

The Recommendation Engine contains content-based filtering technique and collaborative filteringtechnique algorithms that have been combined to recommend postgraduate degree programs to users based on user preferences and user's network.

The improvement of recommendation systems is increased on integrating machine learning techniques with ontology-based recommender systems. The system generalizes user preferences through machine learning techniques. The proposed system overcomes sparse data problem of the traditional content-based recommender using Spreading Activation. Furthermore, in paper [6], authors proposed an ontology-based recommender system with rules generated by data mining techniques. Instructors can use this system to predict learners' progress and performance. The results were very satisfying and verified the generated rules with small deviation error prediction. The combination of ontology-based recommender system with machine learning techniques is a promising approach for improving recommendation accuracy. Therefore, after verifying the approaches of the different systems mentioned in section 2, we chose to implement this combination of techniques and processes.

- Content based recommendation

Content based filtering helps to generate a user (interest) profile vector and postgraduate degree program (course) profile vector. Both vectors include number of features and for each feature weight is assigned. In CBF the similarity between the degree program and user preferences are calculated in the created database. For recommending the most appropriate degree program, the similarity degree ranked form high to low. Traditional CBF method, if there is large number of features in a vector the similarity calculation is very high. In this paper, the problem is overcome by the ontology concepts hierarchy of the postgraduate degree program and user profile.

- Collaborative recommendation

The collaborative filtering method includes the type of finding the similar users respective to the target user to calculate the similarity for K-top nearest neighbors and recommend a postgraduate degree programs which have the highest rate. The main disadvantage of CF is that it needs large number of rates to find the similar users. The users are too lazy to rate a postgraduate study program which recommended for them by the system. Therefore, using an ontology for calculating the user profile information prevented that drawback.

- Hybrid recommendation approach

The recommendation lists taken from CBF and CF are combined by mapping the similar postgraduate degrees in each list and computing the similarity degree for similar postgraduate studies. When calculating the final scoring function, the other factors weights such as university rank, university location will be added to the postgraduate degree program weights in the final list. Finally the top most appropriate postgraduate degree program will display for the user.

4. Experimental Results

The accuracy of the postgraduate degree program ontology is checked by retrieving answers for intelligent questions. For querying reasoner and SPARQL query engines can be used. Before connecting the recommendation engine to the ontology, individuals are added in a manual process. Some answers weren't complete due to lack of data.

Scenario 1: After getting data from a specific user profile as an input to the ontology, checked whether the interest areas of that student can be extracted.

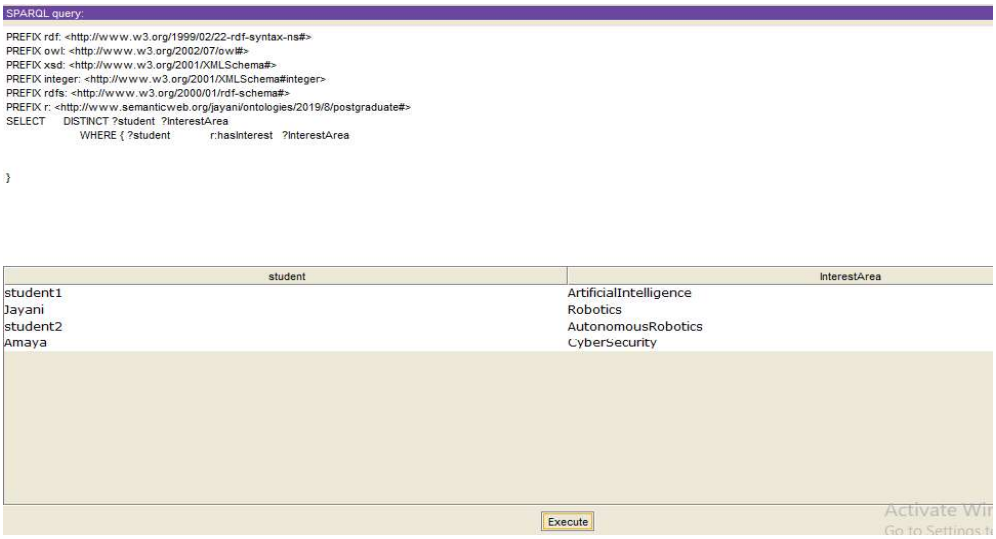

Figure 4: SPARQL Query and output for the interest areas of students.

Scenario 2: When the recommendation system needs to check the current qualification level of a user.

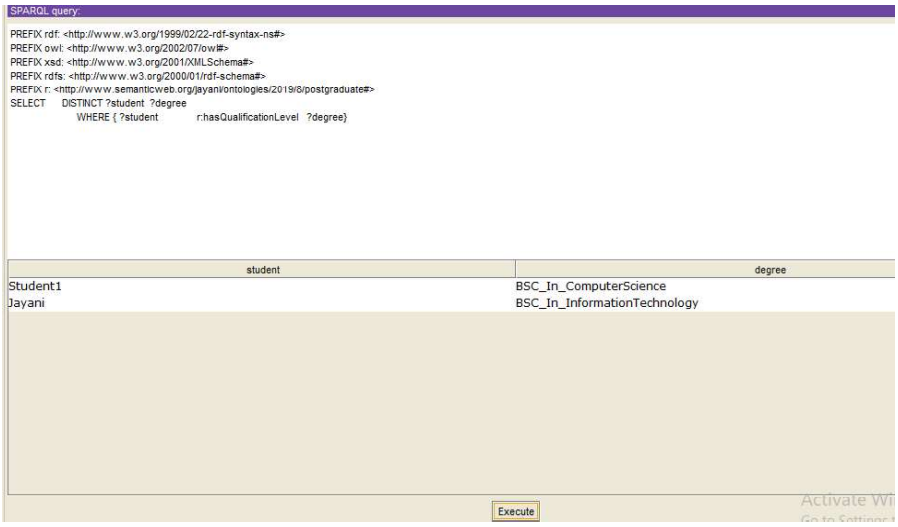

Figure5: SPARQL Query and output for the current qualification level of a student

Scenario 3: The recommendation system wants to check the preferred languages of a user

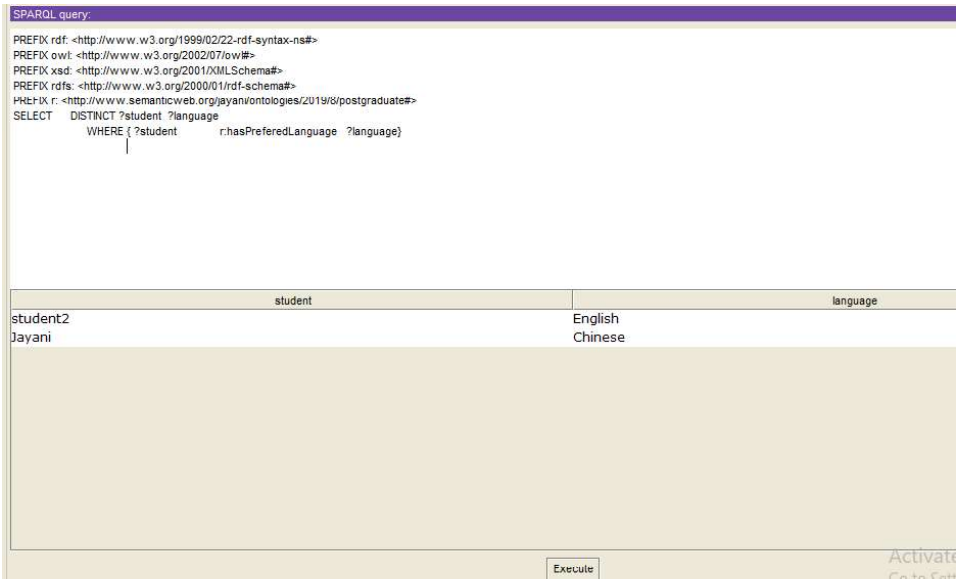

Figure6: SPARQL Query and output for the preferred languages of the degree of a student

Scenario 4: The recommendation engine searches the degree program title which corresponds to the user's interest area.

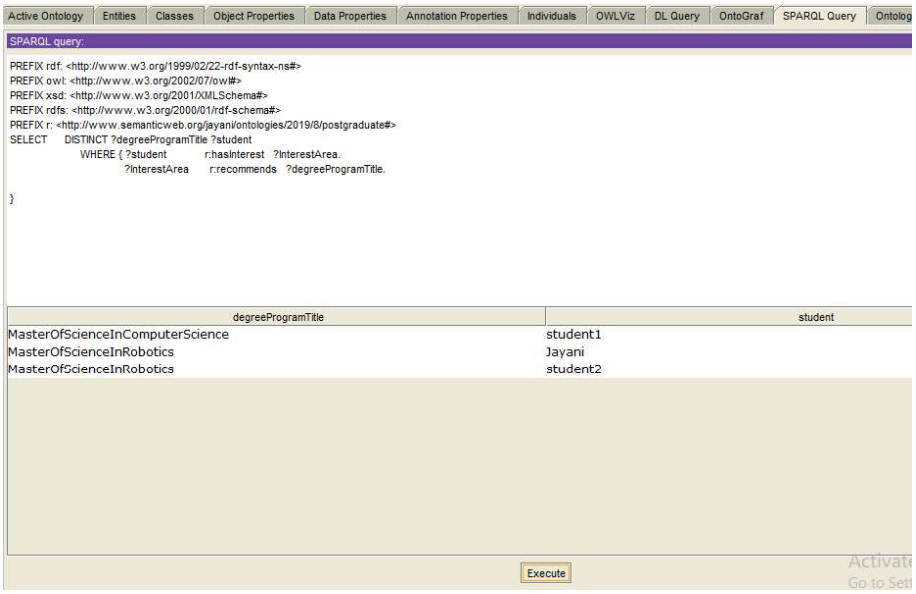

Figure7: SPARQL Query and output of a degree program title of a student which corresponds to the student's preference.

Scenario 5: The recommendation engine extract data from the ontology about the location of the universities which the student's recommended degree titles are matched.

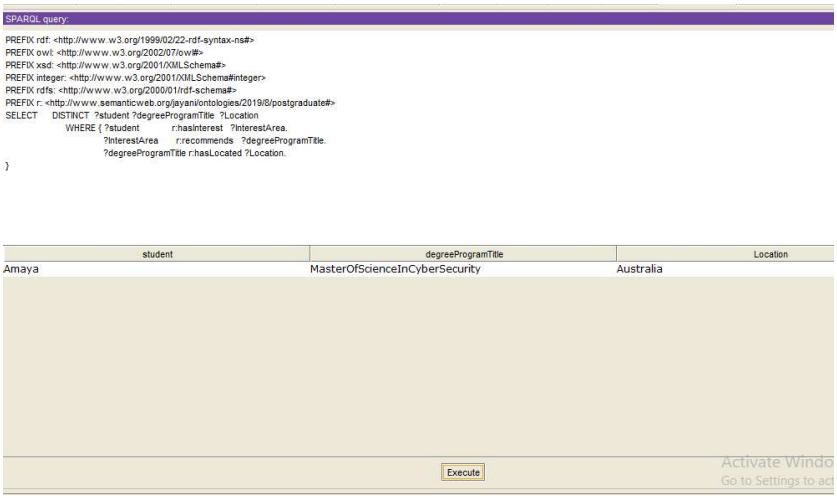

Figure8: SPARQL Query and output of a student for recommending a degree program title which is available in a university corresponds to preference location

5. Performance Evaluation

Precision and recall, F1- measure are the most popular performance evaluation metrics commonly used by information extraction systems. Precision presents a percentage value of number of correctly provided recommendations of the total number of provided recommendations for the users. Following formula is about measuring the precision (See equation 1). (1)

$$\text{Precision} = \frac{|\text{Correctly provided recommendations}|}{|\text{Total recommendations}|}$$

Recall provides the ratio between total number of correctly provided recommendations and number of user profiles. Following formula represents the F1-measure metric (see equation 2).

$$\text{Recall} = \frac{|\text{Correctly provided recommendations}|}{|\text{Total number of userProfiles}|} \quad (2)$$

F1-measure can be used to evaluate the system, other than the mentioned two metrics. Following formula is about measuring F1-measure (see equation 3).

$$\text{F1-measure} = \frac{2\ (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \quad (3)$$

The goal of the evaluation performance was to identify whether ontology based recommendation system perform better than a traditional recommendation system such as TF-IDF. We have developed a test method to evaluate our approach while building a test environment and distributed the survey among 250 final year undergraduates of faculty of Information technology, University of Moratuwa. The calculations are done, using a confusion matrix (true positives, false positives, true negatives, false negatives) respective to retrieving and not retrieving relevant and non- relevant information. Then we compared the precision, recall F1-measure against the considered recommendation systems.

Using the evaluation metrics, we can identify that ontology based recommendation systems can provide more précised recommendations respective to traditional recommendation systems.

6. Conclusions

Selecting a postgraduate degree program is very complicated and tedious process for a student. It is a real challenge to help a student for making a decision for selecting a postgraduate degree program from the available information of digital content. This paper focused on three main aspects.

1. Implement a novel postgraduate degree program that can reduce the problem of information overloading and minimize the available options for selecting a specific postgraduate degree program (personalized recommendation).
2. Present a novel approach for data extraction and data integration from LinkedIn user profiles based on ontology mapping between classes for gaining a comprehensive knowledge about the postgraduate degree programs.
3. Design and develop a hybrid recommendation system that combines the content based and collaborative filtering techniques using the developed postgraduate degree program ontology.

References

1. Mamadou Diaby, Emmanuel Viennet, Taxonomy-based Job Recommender Systems On Facebook and LinkedIn Profiles , 978-1-4799-2393-9/14/$31.00 ©2014 IEEE
2. Mahsa Mehrpoor, Andreas Gjærde, Ole Ivar Sivertsen, Intelligent Services: A Semantic Recommender System for Knowledge Representation in Industry, 978-1-4799-4735-5/14/$31.00 © 2014IEEE
3. CHUNG-YI HUANGl,2, RUNG-CHING CHEN\ LONG-SHENG CHENl,Course-Recommendation System Based On Ontology Proceedings of the 2013 International

Conference on Machine Learning and Cybernetics, Tianjin, 14-17 July, 2013

4. Ján Lang , Tomáš Hnojþík, Information Fragments' Relationships Mining and Their Mapping in Ontologies, SAMI 2016 IEEE 14th International Symposium on Applied Machine Intelligence and Informatics January 21-23, 2016 Herl'any, Slovakia
5. Tyler Morrow, Ali R. Hurson, Sahra Sedigh Sarvestani, A Multi-Stage Approach to Personalized Course Selection and Scheduling, International Conference on Information Reuse and Integration, International Conference on Information Reuse and Integratio,n978-0-7695-6243-8/17 $31.00 © 2017 IEEE DOI 10.1109/IRI.2017.58 2017 IEEE International Conference on Information Reuse and Integration (IRI)
6. Neethukrishnan K V,Swaraj K P , Ontology Based Research Paper Recommendation Using Personal Ontology Similarity Method , 978-1-5090-3239-6/17/$31.00©2017IEEE
7. Bouchra Bouihi, Mohamed Bahaj, An Ontology-based Architecture for Context Recommendation System in E-learning and Mobile-learning Applications, 3rd International Conference on Electrical and Information Technologies ICEIT'2017
8. Kunal Shah, Akshaykumar Salunke, Saurabh Dongare, Kisandas Antala, Recommender Systems: An overview of different approaches to recommendations, 2017 International Conference on Innovations in information Embedded and Communication Systems (ICIIECS)
9. Mohamed Frikha, Mohamed Mihiri, Faiez Gargouri, Using Social Interaction Between Friends in Knowledge-based Personalized Recommendation, 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications
10. Boya Cheng, Yuan Zhang,Dongxin Shi, Ontology-based personalized learning path recommendationfor course learning, 2018 9th International Conference on Information Technology in Medicine and Education
11. Gohar Reham Chughatai, Jia lee, Asif Kabir, Rashid Abbasi, Muhammad Arshad Shezad Hassan, Ontology-Based Academic Article Recommendation, 978-1-7281-1536-8/18/$31.00 ©2018 IEEE
12. S.R.Rimitha, Vedasamhitha Abburu, Annem Kiranmai, K. Chandrasekara, Ontologies to Model User Profiles in Personalized Job Recommendation, 98978-1-5386-5323-4/18/$31.00 ©2018 IEEE
13. Charbel Obeid, Inaya Lahoud, Hicham El Khoury, Pierre-Antoine Champin , Ontology-based Recommender System in Higher Education, Track: The Third Edition of Educational Knowledge Management Workshop WWW 2018, April 23-27, 2018, Lyon, France
14. Ling Zeng, Tonglin Zhu, Xin Ding, Xiaodong Xu, Study on Construction of University Course Ontology, 2009 Fifth International Conference on Semantics, Knowledge and Grid
15. Ayesha Ameen, Khaleel Ur Rahman Khan, B.Padmaja Rani, Creation of Ontology in Education Domain, 2012 IEEE Fourth International Conference on Technology for Educatio