

Tool Based Learning: A Classroom Guide for Information Retrieval

Praveen M.Dhulavvagol¹, Anand S Meti², Sunil V Gurlahosur³, Uday Kulkarni⁴, Somashekhar Patil⁵

^{1,2,3,4,5}B.V.B College of Engineering and Technology
Hubli, India

¹praveen.md@bvb.edu, ²anandsmeti@bvb.edu, ³svgurlahosur@bvb.edu, ⁴uday_kulkarni@bvb.edu, ⁵skpatil@bvb.edu

Abstract: Students today have instant access to information through technology and the web, and they are able to manage their own acquisition of knowledge through informal learning. As a result, traditional teaching and learning methods are becoming less effective at engaging students and motivating them to achieve the desired goals. Today's students are presented with content-centric assignments that meet standards but lack a real-world context and opportunities for active participation. So a step was initiated to enhance the student learning by introducing tool based learning methodology, which provides the platform for better understanding of real world problems and able to develop deeper knowledge of the subject they are studying.

In this paper we discuss how we carried out the process of tool based learning approach for information retrieval(IR) course, This methodology looks at building a prototype for information retrieval applications which confine to recommendations, classifications or clustering based technics and using web as a medium for building up good interactive user friendly applications. This TBL provides a flavour of combining different technics of IR, where students need to have active participation; the overall results show the better performance in their thinking and problem solving skills. The semester end results reveal that they have achieved a good understanding of the concepts, techniques and tools that were used in the course to develop prototype applications.

Keywords: TBL, IR, Prototype, Program outcomes

I. INTRODUCTION

Current information retrieval tools mostly use keyword search, which is unsatisfactory option because of its low precision and recall. In this paper, we consider concept based information retrieval model as a new and promising way of improving search on the web. Informally, concept-based information retrieval is search for information objects based on their meaning rather than on the presence of the keywords in the object. In the last few years, concept-based information retrieval tools have been created and used mostly in academic and industrial research environments

We at the BVB College, Information science department have categorized all the courses into 4 different verticals.

The 4 verticals are as follows:

- Data Engineering
- Network Engineering
- System Engineering
- Software application development

The IR course falls under data engineering vertical, this course is offered as elective for 7th semester, the number of students registered for this course on an average is 60, the prerequisite for the course are Database Management system and Linear Algebra and probability subjects.

The different courses which are offered in this vertical are data structures, database management system, data mining, advance database management system, business intelligence and distributed cloud computing. Students study these courses from 3rd semester to 8th semester, by the time they become graduate they will have a thorough knowledge in data analytics and database technology and the students can shape there career in database technology.

Curriculum Design for IR:

IR is very interesting field were we can learn different retrieval techniques such as supervised and unsupervised learning techniques. IR course can be thought with two aspects considering text data and multimedia data so lot of discussion were carried out, and then it was fixed to teach IR course considering text data, and accordingly the syllabus was designed referring

Praveen M.Dhulavvagol¹

B.V.B College of Engineering and Technology
Hubli, India

¹praveen.md@bvb.edu

various textbooks basically we followed “Manning, Raghavan and Schutze, Introduction to Information Retrieval, text book to teach the students, and many other reference books were referred. Before we discuss individual topics of the course, we will look into the objectives of this course and how students can adopt the concepts and techniques of this course to carry out their research, capstone and REU projects, we can observe that in today’s academic environment, the students learning an IR course will have different motivations, expectations. To simplify things a bit, we will consider the students with respect to their expected working relationship to IR systems i.e. with respect to developing an IR model for a suitable application or doing research or implementation projects in the IR domain.

Considering the above aspects the syllabus for the course was designed with the following topics: Introduction to IR, IR Models, Query Operations, Performance Evaluation, Categorization, Clustering, and Search Engine & Link Analysis. With respect to these topics the course outcomes were designed and are as follows:

Course Outcomes-(CO)

1. Demonstrate the working of information retrieval Systems and the way such systems interact with Users and affect their productivity.
2. Apply problem solving skills on IR Storage Methods, Retrieval Models and Retrieval Utilities.
3. Evaluate IR systems based on relevance of the documents retrieved, using relevance feedback and query expansion techniques.
4. Analyse and interpret various forms of text Categorization and clustering algorithms to increase the efficiency of IR system.
5. Explain the concepts, applications of web search engine and Link analysis.
6. Implement different IR techniques/models for the Chosen dataset.

Once the Course learning outcomes were decided accordingly the program outcomes and the performance indicators were decided to achieve the above course outcomes. The articulation matrix of table 1 shows the mapping of course outcomes with the abet a-k criteria

Table 1: Course Articulation Matrix for IR Course

Course Outcomes-CO	1	2	5	9	10
CO-1	M				
CO-2	H				
CO-3	H	M			
CO-4	M	M		M	
CO-5	H				
CO-6	H		M	H	M

Course Design and Delivery Approach

The course is offered as department elective subject for 2012-13 to 2014-15 batch students. For the 2012-13 batch students the course was delivered with minor 1,2 and

course seminar as activity for 50 marks CIE and Sem end exam for 50 marks so total 100 marks. Feedback from the students were taken and found that students were thorough with theoretical aspects of the course but they were feeling difficulty to apply these concepts to implement IR model or system in their project work, So accordingly in the next academic year 2013-14 there was change in evaluation pattern instead of course seminar, course activity of implementing the IR models/techniques was included so that students should achieve the program outcome element 2, students implemented the IR models/techniques using java/c/c++ platform with a small data set and implemented different IR techniques/methods. Again the feedback was taken and students gave good feedback and told its interesting subject and we got to know how different search engine works and different ways of retrieval methods, the course activity of implementing IR models/techniques helped in their capstone project development.

To make learning more effective a tool based learning approach was introduced for 2014-15 batch students. The CIE activity involved to implement and evaluate different IR techniques/models using open source available tools. Students underwent this activity and we were able to achieve program outcome element 2, 5, 9, 10 with the Degree of compliance L: Low M: Medium H: High. The various program outcomes achieved is as shown in below table 2.

Table 2: Program Outcome elements addressed in the Course and corresponding Performance Indicators

Outcome Element:	1.1 Demonstrate the competence in mathematical modelling. 1.4 Demonstrate competence in Computer Science Engineering
Indicator	1.1.1 Apply the knowledge of discrete structures, linear algebra, statistics and numerical techniques to solve problems.
	1.4.1 Apply theory and principles of computer science engineering to solve an engineering problem
Outcome Element:	2.2 Apply theory and principles of computer science engineering to solve an engineering problem 2.4 Demonstrate an ability to execute a solutions process and analyse results
	2.2.4 Compare and contrast alternative solution/methods to select the best methods
	2.4.2 Analyse and interpret the results using contemporary tools.
	2.4.3 Identify the limitations of the solution and sources/causes.
Outcome Element	5.1 Demonstrate an ability to identify/create modern engineering tools, techniques and resources
Indicator	5.1.2 Create/adapt/modify/extend tools and techniques to solve problems
Outcome Element	9.3 Demonstrate success in a team-based project
Indicator	9.3.1 Present results as a team, with smooth integration of contributions from all individual efforts
Outcome	10.3 Demonstrate the ability to

Each of the performance indicators were mapped to the course outcomes and were measured, the below table 3 shows the mapping of course outcome with the performance indicators.

Table 3: Course outcome mapped to performance indicators

S.N	Course Outcomes	Performance Indicators
1	Demonstrate the working of Information retrieval systems and the way such systems interact with Users and affect their productivity.	1.4.1
2	Apply problem solving skills On IR Storage Methods, Retrieval Models and Retrieval Utilities.	1.1.1
3	Evaluate IR systems based on relevance of the documents retrieved, using relevance feedback and query expansion techniques	1.4.1 2.2.4
4	Analyse and interpret various Forms of text categorization and clustering algorithms to increase the efficiency of IR system.	1.4.1 2.4.2 2.4.3
5	Demonstrate the concepts, applications of web search engines and link analysis	1.4.1
6	Implement different IR Techniques/models for the chosen dataset.	5.1.2 9.3.1 10.3.1

Literature and Teaching Methodology:

The chapters discussed in introduction section of this Paper are explained referring the “Manning, Raghavan and Schutze, text book. The more advanced topics currently discussed in research are addressed during the course activity or seminars/presentations by the students.

Here we discuss three different ways to make the students understand the subject in-depth, in the first approach research paper having relevance to the course were given to the students and were asked to understand the IR technique used, the methodology and the evaluation methods and the students need to prepare a report on that and present the same this activity was conducted for 2012-13 batch students the result in the results section depict that we were able to achieve abet criteria 1,2,9 but we could not achieve some of the PI’s.

So for 2013-14 batch students the CIE activity was modified and was asked to implement the IR techniques/models for a small data set using any platform and technology. By doing this activity we were able to address all the performance indicators which we were unable to achieve for 2012-13 batch students. The feedback

of students was good as almost all agreed that they were able to map the concepts and correlate them and were able to apply the techniques for the chosen dataset so in this way the learning has enhanced.

For 2014-15 batch students we introduced tool based learning approach were they need to use open source tools to develop/implement an IR system for the chosen data set and need to evaluate results and do the comparative study between different techniques. This was additional which helped in getting good score for all the PO’s and PI’s which are mentioned in table 3. This helped students to explore different tools and techniques and they were able to distinguish different techniques and to choose the best technique appropriate for the chosen data set. TBL approach had more effectiveness in student learning as compared to previous batch students.

Tool Based Learning Approach:

In order to design information retrieval (IR) learning environments and its importance in day today life, it is important to explore learning outcomes of different pedagogical solutions. The focus of this study is the assessment of learning outcomes in tool based learning environment compared to traditional teaching methodology. The average 60 students of three academic years were selected for this study, and the analysis illustrates their learning outcomes regarding both conceptual change and development of IR skills. Course activity was conducted to explore different retrieval tools and to measure the performance of different IR technique’s/models choosing appropriate data set.

Students in the tool based learning environment changed their conceptions more regarding aspects of IR and concentrated more on management of search process. Development of information retrieval (IR) systems and applications has made IR a commonplace activity in different work, educational and other activities. End user searchers are now afforded with tools and have access to a wide variety of information sources. Information access skills have gained attention both in professional and all-round education. Different open source tools were used by the students to develop IR applications/systems. So this TBL approach enhanced their learning and concept understanding in-depth and better. Lucene is the most widely used information retrieval library many groups of students have utilized this tool library files to develop their applications, similar to lucene other tools and library files are being used in the course activity process.

IR Course Learning Framework:

The below figure 1 depicts the over view of IR course Learning which involves studying of retrieval models, classification and clustering techniques, the outcome of this learning framework will lead the students to carry out capstone and REU projects and intern will help them in their placements.

Students IR Course learning frame work

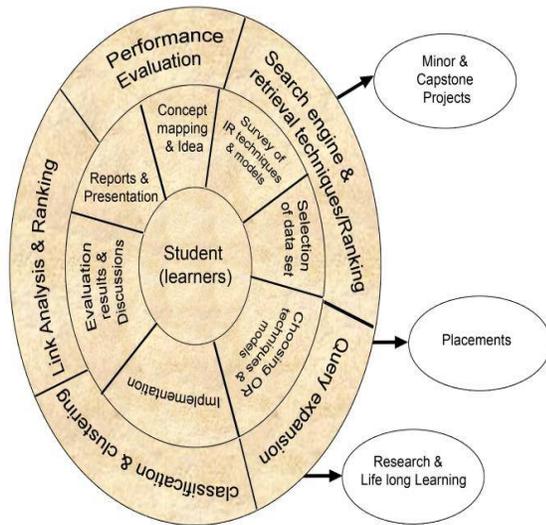


Figure 1: IR Learning Framework

Benefits of TBL Approach:

1. Design and analysis skills improvement
2. In-depth concept understanding
3. Motivated the students to study the subject beyond the Circular
4. Lifelong learning with capstone and REU projects
5. Exposure to open source tools used for developing IR Models
6. Performance evaluation techniques and methods

TABLE 4: IR Course Learning Approach

Features Considered	Case Studies/ Seminars	Projects/ Group Activity
Concept based learning on various IR Application domains.	√	√
Tool based learning approach	√	√
Self-directed learning	√	√
Peer learning(team work)	√	√

The above table 4 provides the different features with the following outcomes:

- Provides a learning framework for better understanding of the subject.
- Understanding and applying of concepts will help in conceptual based learning.
- Explore possible techniques/methods to solve the identified problem.
- Report writing and presentation skills will be improved.

Course Activity Conduction Process:

Activity Conduction Process

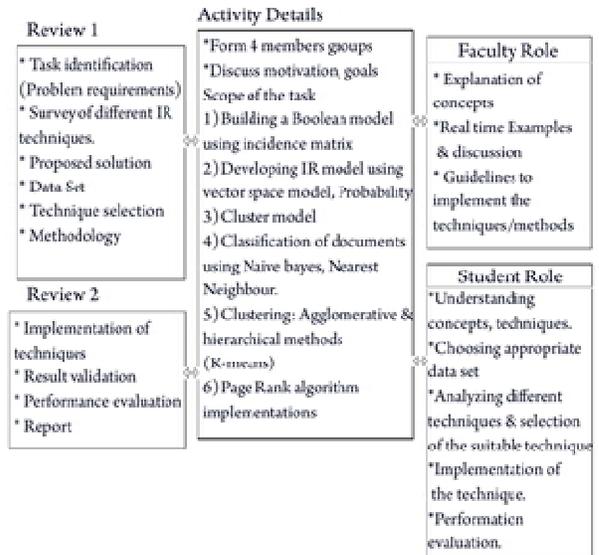


Figure 2: Course Activity Process

Case Studies:

The course activity was to implement IR Techniques/models for the chosen data set the topics chosen by the students are discussed in table 5.

Table 5: Case Studies

Sl.n	Topic Name
1	Implement your own web crawler
2	Implement K-means clustering technique considering your own data set
3	Apply single link method to classify the Documents/data. Assume your own relevant data set
4	Vector Space Model: Implementing a vector space model. Calculate weights for the query and documents and rank the documents. Comparative study: Euclidean distance vs. Cosine Similarity
5	Implement rochho relevance feedback mechanism
6	Implement k-means clustering technique for clustering a set of documents. Choose appropriate document set
7	How page rank algorithm works, work out with simple example and demonstrate the same

Course Activity Rubrics:

Table 6: Evaluation Rubrics

Parameters	Marks (%weightage)
Data set, concepts, techniques and process followed	20%
Implementation	40%
Result evaluation	20%
Report and Presentation	20%

Achievement and Result Analysis:

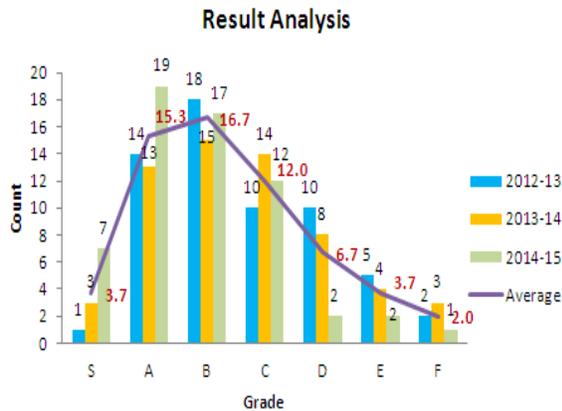


Figure 3: SEE Result Analysis

By introducing tool based learning approach we can observe from figure 3 that the entire outcome elements mentioned in the table 3 have been addressed & the analysis of last three academic years result are shown in the figure 3. From this analysis we observe that there is increase in the S grade and we can see the grade shifting from D to C, C to B, B to A. this intern helped in student learning and were able to map and correlate the concepts. The course activity helped them to apply the concepts/techniques to develop different IR models and systems.

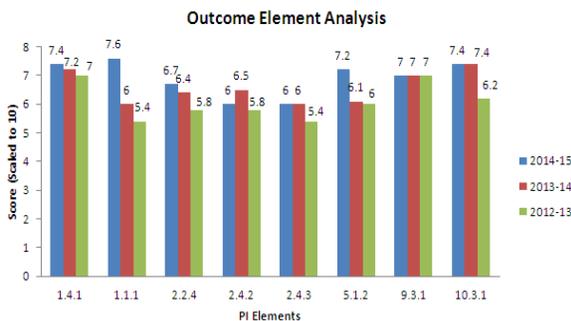


Figure 4: Outcome Attainment (PI's)

The attainment graph shown in figure 4 indicates that there is increase in the attainment of PI's (1.1.1, 2.2.4) compared to last academic years.

Conclusion:

Tool based learning approach helped students to understand the concepts in-depth and they were able to choose the appropriate techniques for the chosen data set, this helped students to compare, contrast different techniques and select the best suited one. This led students to evaluate the results and do performance evaluation Course activity helped Students to develop the applications using different open source tools and they were able to analyse the results. Students were stimulated to carry out capstone and REU projects under the IR domain. Even we observed that they performance was good in semester end Exams compared to previous academic years.

References:

- [1] Allan, J., B. Croft, A. Moffat, M. Sanderson, Frontiers, challenges, and opportunities for information retrieval: Report from SWIRL 2012 the second strategic workshop on information retrieval in Lorne. SIGIR Forum, 2012. 46(1):2-32.
- [2] Meena, M.S.; Anand S Meti; Praveen, M.D. "A Categorized laboratory experiments: Enquiry based approach: A case study" MOOC Innovation and Technology in Education (MITE), 2013 IEEE International Conference in Year:2013 Pages: 265 – 269
- [3] Helen Crabtree "The Impact of Student Centered Learning: An Evaluation of Tutor and Student Experiences", 2006
- [4] Fernandez-Luna, J. M., Huete, J. F., MacFarlane, A. & Efthimiadis, E. N. (2009). Teaching and learning in information retrieval. Information Retrieval, 12(2), pp. 201-226.
- [5] Gordin, D. N., & Pea, R. D. (1995). Prospects for scientific visualization as an educational technology. Journal of the Learning Sciences, 4(3), 249–279.
- [6] Gareth J. F. Jones "An Inquiry-Based Learning Approach to Teaching Information Retrieval" School of Computing, Dublin City University, Dublin 9, Ireland.
- [7] Fernández-Luna JM, Huete JF, MacFarlane A, Efthimiadis EN (2009a) Teaching and learning information retrieval. Inf Retr 12(2):201–226.