

Web Mining: A Key enabler for Distance Education

D.Santhi Jeslet¹ Dr . K.Thangadurai²

GJCST Computing Classification
H.2.8, J.1, H.3.5, H.5.3

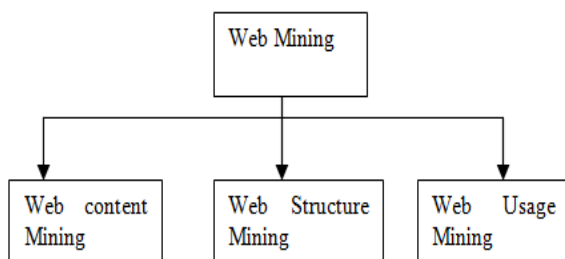
Abstract-This paper deals with introduction of one of the application of data mining which is known as web mining. It discusses about various categories of web mining. It also deals with the application of web mining in distance education and describes the possibilities of application. In this fast world everyone wants to be educated by acquire huge knowledge in a short duration. They do not want to spend some fixed time for their education. Whenever a person is free they can learn and gain the knowledge.

Keywords-Data mining, web mining, distance education

I. INTRODUCTION

Now-a-days many organizations accumulate huge amount of data. This leads to swell the size of the database as the time passes. Traditional database queries access a database using SQL queries. The output of this could be data from database that satisfy the query. This output cannot give any novice information or correlation among the data. So we need a technique that finds the hidden information from data collection in a database community which is of large size. This technique is called the "Data Mining". It discovers valid, novel, potentially useful new correlation and new trends from the large amount of data. Data mining uses pattern recognition techniques, statistical and mathematical techniques for its discovery.

In the recent trend, lots of databases are available in the web. Not only the database, many valuable informations are also available in WWW. So the search area for any information has become very vast. Web mining is an application of data mining which uses the data mining techniques to automatically discover and extract information from Web documents/services. It can also be applied to semi-structured or unstructured data like free-form text. Web mining activities can be divided into three categories: content mining, structure mining, and usage mining. The taxonomy of Web mining is depicted in the figure.



About-¹ Department of Computer Science, M.G.R.College,Hosur,TN, India
(e-mail: santhi.jeslet@rediffmail.com)

About-²Department of Computer Science, Government Arts College(Men)
Krishnagiri- 635 001, India

1. Web Content Mining: It is the process of discovering useful information from the web which may be in the form of text, images, audio and video. For the discovery it uses the techniques of Artificial Intelligence (AI), Database and most specifically Data Mining (DM).
2. Web Structure Mining: It helps to derive knowledge of interconnection of documents, hyperlinks and their relationships. It uses graph theory to analyze the node and connection structuring of a web site.
3. Web Usage Mining: It is also called as web logs mining. This helps to judge about the usage of a web page. It uses computer network concepts, artificial intelligence and database.

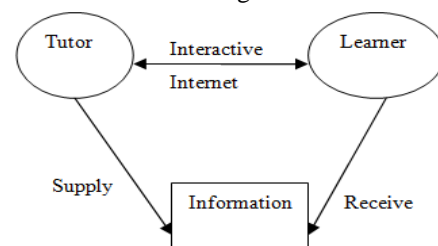
II. OBJECTIVES OF DISTANCE EDUCATION

In the last few decades education has undergone many changes. Class room teaching is needed for face to face education which comprises of class room, presence (physical) of some learners and a teacher/tutor. Here teacher/tutor plays a vital role. But by the introduction of distance education, the interactions between the tutor and the learner have been very much reduced. Even the interaction between the learners has become almost zero.

The main aim of distance education is to make the society to acquire more knowledge irrespective of the place where they are. Those who do not want to stick on to the rules of regular education system, prefer to earn knowledge through distance education. It also encourages working people to attain their learning goals.

III. HOW WEB HELPS IN DISTANCE EDUCATION?

The communication between the tutor and the learner can be enhanced by the introduction of distance education through web. Here learners work individually at their own place, with the help of some study materials i.e. system, computer program and internet. Time and space limitation of education disappears. Tutors interact with the student and the learner interacts with the tutor via internet. The tutor supply information and learner gets it.



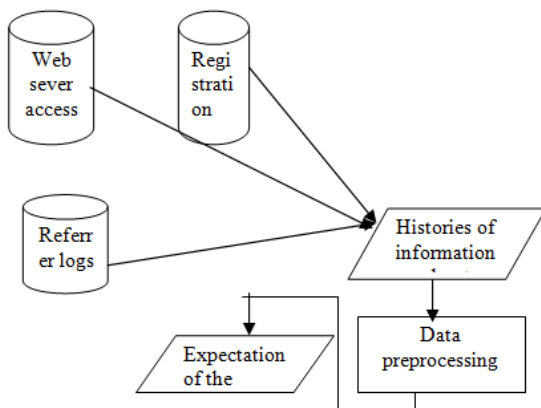
Since many softwares are very simple and user friendly, no need to get special training for working with computer. Power of computers makes student to improve their ability.

Role of tutor is entirely changed. Tutor communicates and leads the course of their learning path. Learners will be grouped. They learn from each other and they also assess each other. It allows the learners to apply their knowledge in different situations and to solve practical problems according to the feedback of their own action. These changes in educational system have developed constructivism. Constructivism means learners involve actively constructing meaningful knowledge through experience.

IV. APPLICATION OF WEB MINING IN DISTANCE EDUCATION

Organization that is responsible for distance education collect huge volume of data, which are generated automatically by web servers and collected in the server access logs. They also collect information from learner (referrer) logs which contain information about the referring pages for each page and also from user registration. Through this an organization can get idea about thinking styles, learns their expectations and also about the web site structure. This helps to improve the efficiency of the web site that is responsible for improving the knowledge of the learners.

Before gathering histories using mining algorithms, number of data preprocessing issues such as data cleaning has to be performed. The major preprocessing task is data cleaning. This is used for removing irrelevant information in the server log.



The extracted access histories of each individual learner are representing the physical layout of web sites with web page and hyperlinks between the pages. Once user access histories have been identified, perform web page traversal path analysis for customized education and web page association for virtual knowledge structures.

By using different path analysis such as graph representation we can determine most frequently traversal patterns form the physical layout of a web site. Path analysis is performed from two points of view: aggregate and individual path. Aggregate path includes the process of clustering the registered learners. The web site database has the registered

learner's details. This can be segmented by one of the clustering techniques to discover learners with similar characteristics. By using this we can determine most frequently visited paths of learners. Individual path helps to determine a set of frequently visited web pages accessed by a learner during their visits to the server.

By discovering such aggregate and individual paths for learner in distance education helps in the development of effective customized education. Associations and correlation among web pages can be discovered using association rules. This guides to discover the correlations among references to various web pages available on the server by a learner or learners. Based on this the tutor can also judge the standard of the learner.

V. CONCLUSION

Web mining in distance education provides a lot of open teaching resources, so that people can teach and learn anytime and anywhere. It helps the organization that is responsible for distance education to discover the learner's access habit and the study interest. It guides the teacher to adjust his/her teaching techniques and the speed of teaching depending on the learner's knowledge. So web mining technology is a key enabler of distance education.

VI. REFERENCES

- 1) Youtian QU, Lili ZHONG, Huilai ZOU, Chanonan WANG. "Research About The Application Of Web Mining In Distance Education Platform. Scalable Computing And Communication, Eighth International Conference On Embedded Computing,2009.SCALCOM EMBEDDED COM'09 International Conference On Digital Object Identifier
- 2) WANG Jian And LI Zhuo-Ling. Research And Realization Of Long-Distance Education Platform Based On Web Mining, Computational Intelligence And Software Engineering 2009, Cise 2009, International Conference On Digital Object Identifier
- 3) Sung Ho Ha, Sung Min Bae, Sang Chan Park. Web Mining For Distance Education, Management Of Innovation And Technology,2000,ICMIT 2000, Proceeding Of The 2000 IEEE Conference On Volume 2, Digital Object Identifier
- 4) Zhang Yuanyuan, Mo Quian. Research Of Constructivism Remote Education Based On Web Mining , Education Technology And Computer Science 2009, ETCS'09, First International Conference On Volume 2, Digital Object Identifier
- 5) Margaret H. Dunham And S. Sridhar. Data Mining: Introductory And Advanced Topics
- 6) Pieter Adriaans And Dolf Zantinge. Data Mining
- 7) Jiawei Han and Micheline Kamber. Data Mining Concepts and Techniques