

## **EDUCATIONAL DATA MINING BASED ON DATA MINING: A CASE STUDY**

**Prantosh Kumar Paul**

FBAS, Bengal Engineering & Science University, Howrah, WB, India

Email: [prantoshkpaul@gmail.com](mailto:prantoshkpaul@gmail.com)

### **ABSTRACT**

*Educational Data Mining is one of the valuable tools for Educational and related Data Collection, processing and further activities is called Educational Data Mining [EDM]. Data Mining is actually a type or retrieval or discovery; it is also a kind of techniques as well as procedure which is dedicated to analyzing and retrieval of Data from various storage system and finalize it in to useful content. Similarly, Web Data Mining is nothing but the Information Analysis and EDM is too. This is a case study paper which is restricted on Web Data Mining; including its basic application and utilization. Though, paper specially illustrated aspects on Education related Data Mining called Educational Data Mining and its several facet.*

**Keywords:** Data Mining, Education Technology, Information Retrieval, Data Engineering, Information Science, Knowledge Discovery, Informatics, Education Informatics

### **INTRODUCTION**

Web Data Mining is the application as well as utilization of Data Mining and similar technique form the web which include data from the web documents, as well as hyperlinks between and among the documents and content usage lot of website and other places[01, 04, 05]. In broader sense, Web Data Mining is including Web content mining, Web Structuring, Structure mining, web usage mining and so on. Similarly Educational Data Mining [EDM] is discovery of knowledge from the educational database collecting, selecting as well as Data and content from the websites related to education and learning is called Educational Data Mining [EDM]. It helps so many activities, but of which study as well as analysis of education data and content treated as most important. It is also able to perform some allied task; out of which analyzing the student capabilities may be treated as important one. Finally, the main aim of Web Data Mining is helpful for so many organization and business houses and individuals.

### **OBJECTIVES**

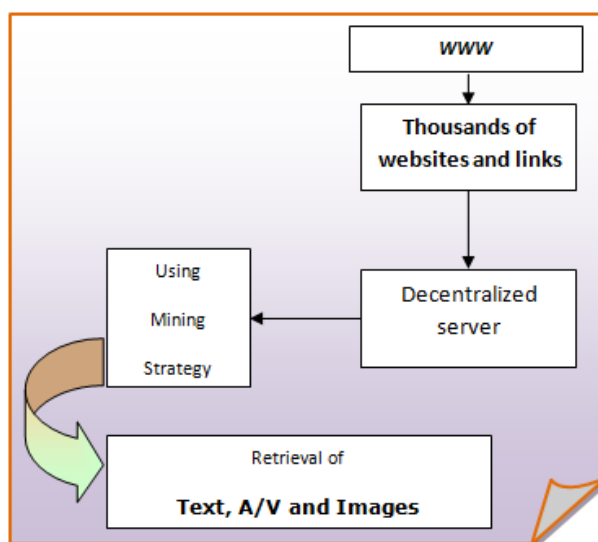
This is a review paper and case study on Web Data Mining and Educational Data Mining [EDM]; paper is mainly dedicated to know about the Data Mining and Web Mining;

including its evolution with special reference to its contemporary benefits; though the paper is also talks about various approaches of Web Data Mining and so on.

Paper tries to illustrate the application of Web Data Mining in so many educational activities. And also highlights on to learn the contemporary and future trends on Web Data Mining and related facet. We try to produce some diagram conceptual on Web Data Mining, Educational Data Mining [EDM] and so on.

## What is Web Data Mining?

Web Data Mining is a technique which is needed to extract from Data from the Web or Internet or from any other similar means. The other online means may include data, full text document and data, hyperlinks from the documents results from the Web. This is related with Webometrics and Web Analysis; Primary Data from the Web may be online and offline. Web Data Mining is including structure mining; including extraction of the structured information from the unstructured Web Data Source. Though depending upon need this situation may be structured in nature. Virtually, Web Data Mining basically helps to see the actual content or data effectively and effectively within a few moments [06, 08, 03].



**Fig: 1:** ABCD of Web and WWW to the masses

## Data Mining and Web Data Mining: Background

The utilization and application of Data Mining is the method of analyzing data for different large data and information sector. During 1990's Web Data Mining become popular in extension data, row data, hyper text, images, records, hyperlinks, tags, http tags, application servers and so on.

During origin or initially, Web Data Mining is treated also treated as extended version and sub system of Data Mining, though today Web Mining is moving to more abstract level where semantics play an important role. The latest of this field is treated as Knowledge Discovery [06, 09, 03].

## Style and Approaches of Web Data Mining

Web Data Mining is categories in to their which includes Web content mining, web structure mining, web usage mining. However, Web Mining as per approach may be classified as two-Agent Based Approach and Database Approach. Agent based Approach is deals with so many tools and techniques powered by Artificial Intelligence and Expert Systems. Some popular types in this category are Intelligent Search Agents, Information Categorization, and Personalized Web Agents. Heterogeneous and semi structured data or Web is fall under Database Approach. This is also called a high level Resources. It is including- Multi Level Database, Web Based Quarry Systems, Hyperlinks, Document Structure and so on [04, 06, 08].

Now Come to the point of types of Web Data Mining; these are as follows:-

**Web Content Mining-** This is the process which is actually the process of extracting knowledge from the content of the Document;

**Web Structured Mining-** Web structured Mining is the process of identification of relationship between Web Pages linked description; so it is also dedicated for the reference and referent to the web;

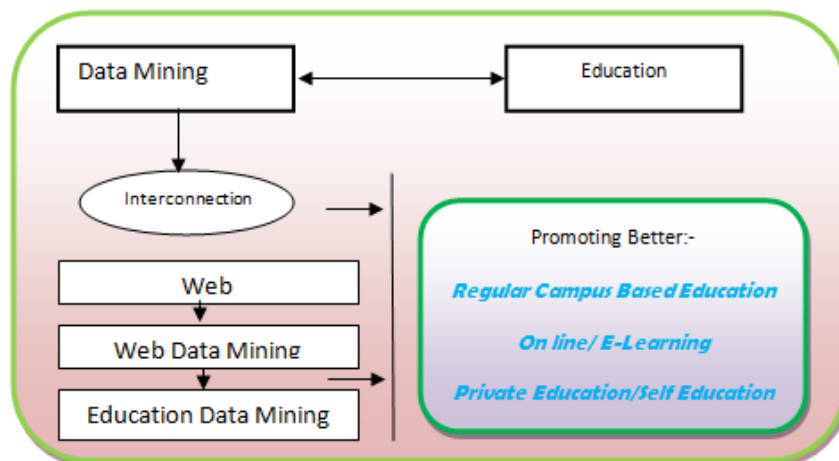
**Web Usage Mining-** This is the lug mining technique needed for processing of extracting interest pattern in Web Access logs.

## Educational Data Mining [EDM]

Educational Data Mining [EDM] may also know as Learning Based Data Mining. This is one of the important domains of Web Data Mining as WDM is applicable in so many domain and dimension. This is actually the discovering of knowledge from the educational databases. This s important and useful the following reason:-

- It is useful for better understanding of the students and the setting where they are basically learned;
- This is also important tool as it helps in the higher education sector in so many activities and particularly discovering improved version of models of the educational knowledge structure;
- Useful in analysis of educational data and content and ultimately for academic Decision Support Systems;
- Clustering of students on a particular point and or parameters[12];
- Designing and Building of Intelligent Course models and full-fledged model of Degree and PG Degree;
- Student activities such as their monitoring, attendance report generation, detecting students relationship and extracurricular activities of the Institutions and so on;
- Educational Data Mining [EDM] is helpful for capabilities of Educational Institutions and so on;

Helps in better and improved pedagogical support designing better module and also useful for come and programme related Information Collection [13].



**Fig: 2:** From Data Mining to Education Data Mining

## Educational Data Mining [EDM]: Features and Methods

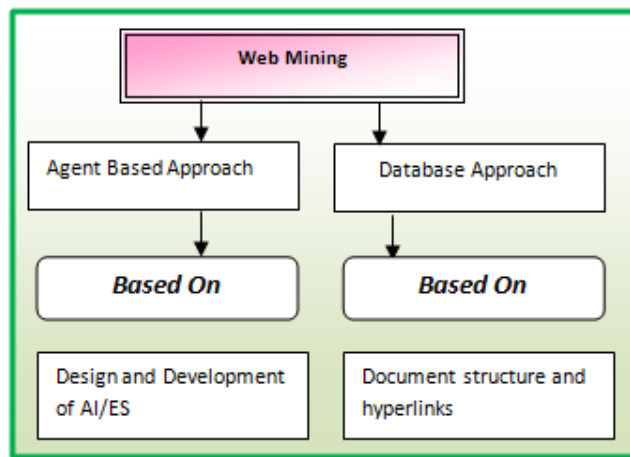
There are so many methods are available for Educational Data Mining; but out of which Baker and Vacet's [2009] methods are important:-

- Out of so many methods, **Prediction** is one of the important and it is useful to find out points that can naturally generate characteristics of data; the prediction is includes classification, regression, density estimation[14, 15,28];
- Discovery and Find out of relationship between variable in a dataset is called **Relationship Mining**. The variable is includes as follows:-
  - Association Value Pattern Mining;
  - Connection Mining;
  - Sequential Pattern Mining;
  - Casual Data Mining and some minor types.
- Apart from Prediction and Relationship Mining, **Model Development** is another important methods and applied in subsequent data as a component for analysis;
- **Making Meaningful Data**- This is the digitalization of data for human judgment.

## Web Data Ming, Educational Data Mining [EDM]: Current Importance and Future Potentials

Web Data Mining is one of the important tools in Data Science and Engineering. It can save the time of the user and also helpful for pre-processing of Information. It keeps healthy Information Division in the organization and helps in exact Information from the Information explosion and overload [05, 06]. Information Explosion is one of the important name now-days. Information is increasing day by day. Today it is very tough to find out

information from the large database and from million and million data and knowledge facet. This is the most important solution of knowledge discovery.



**Fig: 3:** Web Mining and its approaches

Virtually, the use of Web Data Mining is increasing in all most all sector and organization regardless of organization type. Today several software and intelligent systems are emerging and for better information management we need much more intelligent software and information management. The future of Information System will be much more optimization bases. Today, several things are emerging in Educational Data Mining [EDM] such as DCS, courseware generation, MATLAB and so on. Apart from NCR, Clustering, Analyzing student performance and so on.

## FINDINGS

- For sophisticated Web Data Mining we need a special user profile creation for the prompt and advance data output, Web Data networks and so on;
- Web Data Mining is the extension of data mining and also helps its has so many dimension out of which Educational Data Mining [EDM] in popular in academics and similar domain.

## SUGGESTION

- Modern Organization and corporate house need to follow Data Mining tools for healthy Information Management and Decision Making;
- Better and proper content Mining and relationship Management is always needed;
- Universities, Academician, students and other higher educational institutes need to follow and use of Educational Data Mining [EDM] in so many educational activities.

## CONCLUSION

Information is one of the important tool and facet for development. Application and use of information is increasing day by day; similarly several tools and technologies are also emerging out of which Educational Data Mining [EDM] is gaining popularity [04, 07, 09].

The different kind of web data mining is increasing rapidly as it fulfils several information activities and providing right information to right one.

## REFERENCES

1. Cohen, E. B. (2004). Applying the Informing Science Framework to Higher Education: Knowledge Development, Management, and Dissemination. Konferencja Pozyskiwanie wiedzy i zarządzanie wiedzą (Proceedings of the Knowledge Acquisition and Management Conference) May 13-15, 2004 Kule, Poland.
2. Cohen, Eli B. and Nycz Malgorzata (2006). Learning Objects and E-Learning: an Informing Science Perspective. Interdisciplinary Journal of Knowledge and Learning Objects Volume 2, 2006
3. Martin, S.B. (1998). Information technology, employment, and the information sector: Trends in information employment 1970–1995. Journal of the American Society for Information Science, 49(12), 1053–1069.
4. Michael Buckland and Ziming liu (1995).History of information science. Annual Review of Information Science and Technology vol. 30: 385-416.
5. Paul,Prantosh Kumar,R Bhatnagar, Sarmistha Chowdhury “Information Infrastructure: Emphasizing its role, challenges and issues in Higher Secondary School In Indian Context” In ICSSR Proceedings of National Conference on Universalization of Secondary Education: Prospects and Challenges, Dec, 2012, ASE, Amrita University, Mysore Campus, Page-450-457, ISBN-978-81-924422-1-1
6. Paul, Prantosh Kumar, Dipak Chaterjee and Bhaskar Karn ‘Information Management: emphasizing traditional and technology focused approach – An Overview’ in IEEE/CSI/AICTE co sponsored Proceedings National Conference on Paradigm shift in Education Technology & Content Management [ETOM-12]DIT,Techno India, TI University, Kolkata, 02-03/03/2012 Page 58-62.ISBN-978-81-9230347-5, SPS Education Publication, Kolkata,India
7. Prantosh Kr. Pau1, K V Sridevi, Minakshi Ghosh, Ashwina Lama “Education Informatics: Powered By Cloud Computing - the sophisticated Educational Information Infrastructure Building” International Journal of Computer Mathematical Sciences and Applications,ISSN-0973-6786, Page-95-102, Vol. 6, No. 3-4, July-Dec. 2012, SP,New Delhi, EIC-G Ganesan, A N University, AP, India
8. Prantosh Kr. Pau1, H A Mangla, B B Sarangi, Ramanna Chetri, K S Shivraj Roshan Rai “Education Technology: Emphasizing EduNxt Knowledge Transformation Systems of Sikkim Manipal University (SMU), Gangtok, Sikkim, India” International Journal of Embedded Systems and Computer Engineering, July-Dec. 2012, Page-109-113, Research Science Press, India.ISSN-0975-4482, Chief Editor-S Ganesan, Oakland University, USA
9. Paul, Prantosh Kumar,Mrinal Kanti Ghose, Dipak Chaterjee ‘Education Technology: its benefits and utilization with special reference to EduNxt, Knowledge Delivery Model of Sikkim Manipal University-A Study’ in IEEE/CSI/AICTE co sponsored Proceedings National Conference on Paradigm shift in Education Technology & Content Management [ETOM-12]DIT,Techno India, TI University, Kolkata, 02-03/03/2012 Page 148-153 .ISBN-978-81-9230347-5, SPS Education Publication, Kolkata,India

10. Prantosh Kr. Paul, K V Sridevi, Minakshi Ghosh, Ashwina Lama “Education Technology: The Transparent Knowledge Delivery through QPN and Cloud Computing” IJSD-An International Journal, Vol. 12, No. 2, December-2012, Page-455-462 , ISSN-0972-3692, SR, NewDelhi, EIC-Vijay Kumar S, Bangalore
11. Prantosh Kr. Paul, K L Dangwal, Asok Kumar Garg “ Education Technology and Sophisticated Knowledge Delivery” Techno-Learn-International Journal of Education Technology, ND Publisher, New Delhi, Vol. 2, No. 2, Page-169-175 ISSN-2231-4105
12. Prantosh Kr. Paul, K L Dangwal and Dipak Chaterjee, “Information Technology and Advance Computing and their interaction for healthy Education, Techning, and learning: The IKM Approach” Asian Journal of Natural and Applied Sciences,ISSN-2186-8468, Page- 70-77 V-1, No. 4, December-2012, Leena and Luna International, Oyama, Japan
13. Paul, Prantosh Kumar, KL Dangwal “Web Data Mining: Contemporary and Future Trends: Emphasizing Educational Data Mining [EDM]” in EduQuest-An International Journal, ISSN- 0976-7258, July-Dec, 2012, New Delhi Publisher, New Delhi.
14. Prantosh Kumar Paul, Dipak Chaterjee, Ashok Kumar “E Learning: New Age Knowledge Model Delivery through Advance Information Technology and Cloud Computing: An Overview” BRICS International Journal of Educational Research, Vol. 3 No. 1, ISSN-2231-5829, Page-22-25, MM University, Ambala, Haryana, India [ Indexed in Index Copernicus, and others]
15. Paul, Prantosh Kumar, KL Dangwal, M Ghosh “Information Systems & Networks (ISN): Increasing Tools and Technological Application with Special Reference to Role as an Educational Tools” IJCSES, Vol.7 No.2 April 2013, pp.99-103
16. Reichman, F. (1961). Notched Cards. In R. Shaw (Ed.), *The state of the library art*04(01), pp. 11–55). New Brunswick, NJ: Rutgers, The State University, Graduate School of Library Service.
17. Saracevic, T. (1996). Relevance reconsidered. *Information science: Integration in perspectives*. In *Proceedings of the Second Conference on Conceptions of Library and Information Science* (pp. 201–218), Copenhagen, Denmark: Royal School of Library and Information Science.
18. Saracevic, T. (1975). Relevance: A review of and a framework for the thinking on the notion in information science. *Journal of the American Society of Information Science*, 26(6), 321–343.
19. Saracevic, T. (1979a). An essay on the past and future of information science education. I. Historical overview. *Information Processing & Management*, 15(1), 1–15.
20. Saracevic, T. (1979b). An essay on the past and future of information science education. II. Unresolved problems of ‘externalities’ of education *Information Processing & Management*, 15(4), 291–301.
21. Vakkari, S.P. (1996). Library and information science: Content and scope. In J. Olaisen, E. Munch-Petersen, & P. Wilson (Eds.), *Information science: From development of the discipline to social interaction*. Oslo, Norway: Scandinavian University Press.
22. Vickery, B.C., & Vickery, A. (1987). *Information science in theory and practice*. London: Butterworths.

## **ABHINAV**

**NATIONAL MONTHLY REFEREED JOURNAL OF RESEARCH IN ARTS & EDUCATION**

**[www.abhinavjournal.com](http://www.abhinavjournal.com)**

23. Wersig, G., & Neveling, U. (1975). The phenomena of interest to information science. *Information Scientist*, 9, 127–140.
24. White, H.D., & McCain, K.W. (1997). Visualization of literatures. *Annual Review of Information Science and Technology*, 32, 99–168.
25. [www.en.wikipedia.org](http://www.en.wikipedia.org)
26. [www.infosci.cornell.edu/](http://www.infosci.cornell.edu/)
27. [www.ischools.org](http://www.ischools.org)
28. <http://www.libsci.sc.edu/bob/istchron/iscnet/ischron.html>