

# Studying Data Mining and Data Warehousing with Different E-Learning System

Dr. Mohamed F. AlAjmi,  
PhD Head of Quality and  
E-Learning units  
King Saud University  
Riyadh, Saudi Arabia

Shakir Khan  
(Corresponding Author)  
Researcher at King Saud  
University, Riyadh Saudi Arabia  
Nationality Indian

Dr. Arun Sharma  
Head, Department of Computer  
Science Krishna Institute of  
Engineering and Technology,  
Ghaziabad-201206, INDIA

**Abstract**----Data Mining and Data Warehousing are two most significant techniques for pattern detection and concentrated data management in present technology. ELearning is one of the most important applications of data mining. The foremost idea is to provide a proposal for a practical model and architecture. The standards and system structural design are analyzed here. This paper provides importance to the combination of Web Services on the e-Learning application domain, because Web Service is the most complex choice for distance education during these days. The process of e-Learning can be promising more efficiently by utilizing of Web usage mining. Mor07/e sophisticated tools are developed for internet customer's behaviour to boost sales and profit, but no such tools are developed to recognize learner's performance in e-Learning. In this paper, some data mining techniques are examined that could be used to improve web-based learning environments.

**Keywords**— Data Mining; Data Warehousing; e-Learning; Moodle; LMS; LCMS.

## I. INTRODUCTION

Usually the decision-making data are stored in files and databases. The results getting by huge amount of data are not easy, for which the data mining techniques are very constructive. Data mining is the process of taking out information in terms of patterns or set of laws (e.g. association rules, sequential patterns, classification trees) from huge databases. So, it is also known as data or knowledge discovery.

For example, by pulling out demographic data of students' enrolments, the university, college or any institute could get better the qualitative explanation (e.g. information for past's students) of database. Any association does not deal with a single database, but deals with various kind of database means multiple databases but there is the need for fast processing, and integrating of these databases which can be possible by data warehouse. Centralizing data management and revival is often distinct as data warehousing. This centralizing helps the user to maximize access to the data and analyzing it.

The data warehouse supports different types of analyses, including elaborate queries on large amounts of data that may require extensive searching. When databases are set up for queries on daily transactions, they are called "operational data stores" rather than data warehouse. So, a data warehouse is a

storehouse of an organization's electronically stored data [3]. The mechanisms of data warehouse are: retrieval, extract, analysis, transform, load data and managing data dictionary. Data mining, data warehousing, and Online Analytical Processing (OLAP) together form the functionality of decision making or Decision Support System (DSS). The various areas Eof application of data mining and data warehousing are e-commerce, e governance, online shopping, digital library, online reading, e-learning or e-education, etc. Among these, these days e learning is an important application of data mining.

E-Learning is sometimes known as electronic learning or e-learning in which there is no face-to-face interaction between the teacher and the students. Rather than it is web-based learning. It uses Web or Internet technology and delivers digital contents, provides learner oriented environment for teachers and students [4]. So, the environment is not teacher-centric. It may include all types of Technology Enhanced Learning (TEL), where technology is used to support the learning process [5].

For example, in companies, e-Learning is used to deliver training courses to employees and in universities, e-Learning is used for enrolment of students in different courses, provides teaching without any face-to-face interaction, or on-campus facilities, but through internet that is online. As a whole, e-Learning includes Distance Learning (DL), Computer Based Teaching (CBT), Computer Aided Instruction (CAI), and Life Long Learning (LLL) principle. So, we see that, e-Learning consists of various types of databases, storing information for user access. To implement e-Learning, data mining can help to construct e-textbook, e-reading, digital libraries, etc.

Further scope of e-Learning is blended e-Learning which is a combination of face-to-face interaction and online learning. It incorporates online lectures, tutorials, performance and decision support systems, simulations and games, and more [5].

## II. E-LEARNING ARCHITECTURE OR DESIGN

### A. Functional Model

The practical model of an e-Learning structure creates an interface between the mechanisms and the objects of the e-Learning system. It is shown in "Fig. 1".



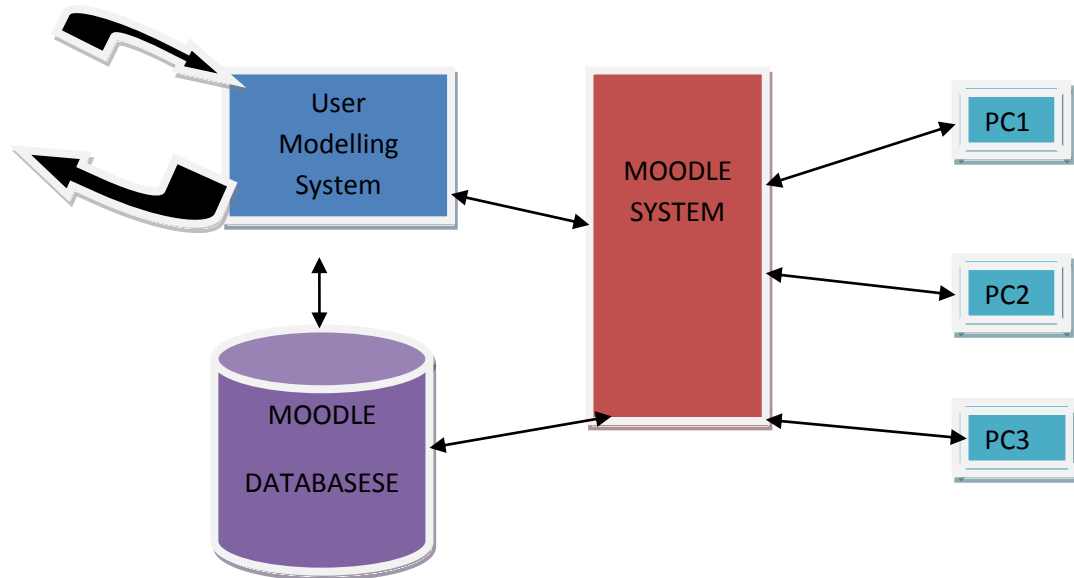


Fig. 1 Moodle Based E-Learning Architecture

The structural design of e-Learning till now does not provide any apparent picture of the e-Learning components. The e-Learning structural design contains two models: the information model and the component model. These two replicas are to be joined and an interface must be defined to attain interoperability. This structural design of e-Learning gives a practical model of the components of e-Learning for the consistency of e-Learning development. The Advanced Distributed Learning (ADL)'s Sharable Content Object Reference Model (SCORM) practical model explains the swap of data within a Learning Content Management System (LCMS) or a Learning Management System (LMS) to track user's progress. But the functionality is not explained by SCORM.

A multi-user atmosphere in which the knowledge developer can create, reuse, manage, store, and distribute digital learning content from a central storehouse is known as LCMS. Here the processes adjoining the learning are managed by LMS. LCMS permits the users to generate and to use again small units of digital instructional learning material.

The incorporated use of metadata arrangements and learning object import and export formats also allows learning objects to be created and shared by multiple tools and repositories. LCMS integrates specifications of metadata, content wrapping, and content communication. The components of LCMS are shown in "Fig 1".

LMS needs the interchange of customer profile and customer registration information with other systems. The position of the course choice and the learner action are offered by the LCMS. The mechanisms and information needed are shown in "fig 1". So, there is an incorporation of LMS and LCMS.

Secondly, the SCORM is developed by US Department of Defense's ADL. This is an "application profile" consisting of a set of terms and conditions. The three main mechanisms of SCORM are:

- 1) *Runtime Environment: The runtime environment is an API describes the interface between learning object and LMS or LCMS to track learner's progress;*
- 2) *Meta-Data: A set of data elements to explain learning contents so that it can simply explore for identified and accessed [7];*
- 3) *Content wrapping: Content wrapping is the release and exchange of structured content i.e. learning objects and courses between different LMS and LCMSs;*

As a course is separated into lessons, and sometimes the lessons are divided into topics The SCORM condition explains two hierarchical levels:

- 1) *Content aggregation: A group of learning resources to construct complex structures, contents aggregations may be nested and may have lower-level blocks of contents which outline a content aggregation;*
- 2) *Resources: Two major types of educating resources are there: SCO and ASSET;*

The stage at which student interacts with the learning content and also the LMS tracks the results is known as SCO. Basically, it is a learning object.

A part of content in form of movie, sound, graphic or other media item is referred as an ASSET. Most ASSETS are started by SCOs as part of their in-house content (e.g. graphics come into view on an HTML page).

### III. STANDARDS IN E-LEARNING

Standards in e-Learning give standardized data structures and communication protocols for e-Learning objects and cross-system workflows [1]. The standards are of the following types:

- 1) *Metadata: Metadata refers to the labelling of learning contents and catalogs to maintain indexing, storage,*

*detection (searching), recovery of learning objects by several repositories of data mining and data warehousing techniques. The data utilized here is known as metadata;*

- 2) *Content Wrapping: Content Wrapping permits the transport of course content from one learning management system to another learning management system. The most significant content wrapping system these days is, ADL's SCORM [7]. The facts of the contents are stored in various databases which can be developed and received by data mining and data warehousing techniques;*
- 3) *User Profile: User Profile consist personal data, learning history, prerequisites, learning plans, degrees and certifications, evaluation of information and contribution status in existing learning;*
- 4) *Student Registration: Student Registration identifies the availability of courses for the learner, also, information about other members of the course.*
- 5) *Content Communication: It gives an interface between student data and previous activity after content is started. The message is developed by ADL's SCORM Object Reference Model.*

This architecture explains the fundamental thought of scattered e-Learning system means the communication of messages through the communication of web service agents, present in each system. Service Provider is the podium that hosts right to use to the service. It is the server in a client-service environment. Service Requester is the function that is looking for and calling upon or initiating the communication with a service. Discovery Agency is a searchable set of service explanation where service providers issue their service descriptions.

According to Xiaofei Liu, Abdulmotaleb El Saddik and Nicolas D. Georganas [1], the discovery agency may be centralized or distributed. Information presented by XML concerning learning is wrapped with the Simple Object Access Protocol (SOAP) arrangement and is swapped between requester and provider. A Web Services Description Language (WSDL) file holding the explanation of the message and information regarding end point is published by the provider to permit requester to create the SOAP message and transmit it to the exact destination.

#### IV. BENEFITS OF DATA MINING IN E-LEARNING

There are several web usage tools to carry out data mining and data ware housing tasks. For, instance, Two data mining and data ware housing tools are WebSIFT and WebLogMiner for pattern detection from web logs [10][11] but these tools are not initiated in e-Learning environment till now because if the educator does not have sufficient knowledge in data mining, can't use these tools to get better efficiency of e-Learning. Web usage mining is a new system, devoted for e-Learning is being industrialized to permit the educators for on-line assess activities [9]. It facilitates the educator to follow the activities in the course web site and take out patterns and behaviours, get better or adapt the course content. For example, one could recognize the paths regularly or frequently

visited, the paths never visited, etc. By analyzing these general traversal paths of the course content web pages or recurrent changes in individual traversal paths, the design of the course can be known to be better fit the requirements of students.

Two types of data mining techniques are used in e-Learning: off-line web usage mining and integrated web usage mining. Off-line Web Usage Mining: Off-line web usage mining is the detection of patterns with a separate application. This pattern detection process permits educators to evaluate the access behaviours, legalization of the learning modules, assessment of the learner's activities, assessment between learners and their access pattern, etc [9]. The model of off-line web mining is a tool for the instructors to apply sequential analysis, association rules and clustering for the detection of relations between the learning actions of learners, interesting prototype of on-line actions and to group parallel access behaviour respectively. So, in off-line web usage mining, incorporated educators can place questions and authenticate the learning models, they utilize as well as the structure of the web site as it is read thoroughly by the learner. It is being observed that off-line web usage mining is a parametric move towards where the parameters are the instructors, educators, learners, etc.

Integrated Web Usage Mining: Contrasting to off-line web usage mining, incorporated web usage mining is the procedure of determining patterns incorporated with e-Learning application. This covers adaptive websites, personalization of actions. Also, suggestion of actions to learners according to their favourites along with their history of actions is done by automatic recommenders in incorporated web usage mining. A recommender-based association rule mining is being expanded currently that consists of facts of finding out applicable association between learning performance and creating association rules are recommended to the learner as the suggested next step in the learning session [10]. So, incorporated web usage mining is a non-parametric approach.

#### V. CONCLUSION AND FUTURE WORK

In this paper, an obvious analysis of the content state of e-Learning standard is being explained. Also, a functional model of dissimilar learning objects is presented here. The swapping of system workflows is also being explained in this paper. E-learning standard gives interoperability between learning systems and tools from several vendors. A standard means of message is set up between dissimilar software applications. This communication is likely by the Web-Services technology.

The Web usage mining technique is explained in this paper, which is a non-trivial procedure of taking out helpful and previously unknown blueprints from the use of Web. The data mining techniques to improve e-education are explained in this paper. Since e-Learning process is a endlessly changeable process, the safety services, the encryption of messages, and the general facts to explain services and services access points in e-Learning systems environments are in call for thought.

Though, several tools using data mining techniques to aid e-Learning system are being developed, the research is still in

progress, since the data record given by the Web Servers are inadequate, so there is a call for more specialized logs from the application side to improve the already logged information.

#### ACKNOWLEDGEMENTS

The authors acknowledge the continuous support of deanship of research in King Saud University since this research is funded by a grant from the deanship under the number of RGP – VPP - 150.

#### REFERENCES

- [1] Xiaofei Liu, Abdulmotaleb El Saddik and Nicolas D.Georganas"AN IMPLEMENTABLE ARCHITECTURE OF AN E-LEARNING SYSTEMS". CCECE 2003 – CCGEI 2003 Montreal May/mai 2003.
- [2] FUNDAMENTALS OF DATABASE SYSTEMS, Fourth Edition, Elmasri and Navathe.
- [3] Data Mining: What is Data Mining, Web site at <http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining.htm>.
- [4] "Introduction to E-Learning "Website at <http://www.chengzhi.net/english/index.htm>
- [5] "E-Learning From Wikipedia to free encyclopedia", Web site at <http://en.wikipedia.org/wiki/E-Learning>.
- [6] C. Romero and S. Ventura , "Data Mining in E-Learning" (WIT Press ,2006).
- [7] Leopold Kause, Carol Fallon "Creating E-Learning Content in Authorware 7 for SCORM1.2 Compliant LMSs and LCMSs", Web site at <http://adobe.com/resources/elearning/article/10packager01>.
- [8] "IMS Global Learning Consortium", Web site at <http://www.imsproject.org/>
- [9] Osmar R. Zaiane, "Web Usage Mining for a Better Web-Based Learning Environment".
- [10] R. Cooley, B.Mobasher, J.Srivastva, "Web Mining: Information and Pattern Discovery on the World Wide Web", Proceedings of the ninth IEEE international conference on Tools with AI, 1997.
- [11] O.R. Zaiane, M.Xin, J. Han, Discovering Web Access Patterns and Trends by Applying OLAP and Data Mining Technology on Web Logs, Proceedings from the ADL'98 –Advances in Digital Libraries, Santa Barbara, 1998.

- [12] E-Learning: A Milestone in the Research of Data Mining; by Sabyasachi Pattnaik, Jui Pattnayak, Priyaranjan Dash [http://interscience.in/SpIss\\_ijcct\\_icct2010vol2\\_no234/25\\_ML](http://interscience.in/SpIss_ijcct_icct2010vol2_no234/25_ML)

#### AUTHORS PROFILE

**Dr Mohammed Fahad AlAjmi** was born in Kingdom of Saudi Arabia. He received his Ph.D in Pharmacy from King Saud University, Riyadh, Saudi Arabia in 2007.He chaired many position in the university and currently working as vice dean for quality and development in Prince Sultan College for EMS affiliated to King Saud University. To date he taught many pharmacy students, more than 30 courses. Students' level varies from primary to undergraduate levels.

**Shakir Khan** was born on 5th Feb, 1978 at Kallanheri in Saharanpur district UP, India. He is working as a Researcher in College of Electronic Learning in King Saud University, Kingdom of Saudi Arabia. He received his Master of Science in Computer Science from Jamia Hamdard (Hamdard University), New Delhi, India in the year 2005, and PhD computer Science, C.S.J.M University Kanpur India. He is member of IEEE. He has actively attended many international conferences and published various research papers in National and International conferences as well as journals. His current areas of interests are in Cloud Computing, Software Engineering, Data Mining and E Learning. Apart from that he worked in the field of Software Development in different MNC companies at Noida India.

**Dr. Arun Sharma**, alumni of IIT Roorkee and Thapar University, received his M.Tech. (Computer Science and Engineering) from Punjabi University, Patiala, INDIA and Ph.D. (Computer Science) Thapar University, Patiala, INDIA. Currently, he is working as head of the department of Computer Science and Engineering Department in KIET school of engineering and Technology at Ghaziabad, India. His areas of interests include Software Engineering, Soft Computing and Database Systems. He has published a number of papers in international Journals and Conferences including IEEE, ACM, Springer, WILEY and others in India and abroad. Dr. Sharma is an active member of IEEE, ACM and Computer Society of India. He is also a member of Board of Studies (BoS) of Mahamaya Technical University (MTU), Noida. He is also on the panel of subject experts and examination for various Universities like IGNOU, BBA University (Central), Lucknow, GGSIP University, Delhi, Thapar University, Patiala, and others. He is also an active member of Editorial Board and Review Committee of several Journals including Journal of Computer Science (USA), International Journal of Computer Science and Security (Malaysia), Research Journal of Information Technology, USA and others.