

Educational Data Mining and Big Data Framework for e-Learning Environment

Prakash Kumar Udupi¹, Nisha Sharma², S K Jha³

¹Department of Computing, Middle East College, Muscat, Oman
prakash@mec.edu.om

²NIMT, Kurukshetra, Haryana, India
nishajan6@gmail.com

³AIIT, Amity University, Noida, India
skjha2@amity.edu

Abstract: E-learning data consists of large volume of educational data and available with complex and hybrid data architecture. Capturing of student performances, student evaluation and student's interaction information are one of the challenges faced by the e-learning software users at the time of analysis. Integrating student data along with educational data for analysis needs complex system design framework. New innovations in e-learning also facilitates augmented learning, adaptive learning, web based learning, activity based learning, and project based learning. Education technology interventions using learning management system, content management system, advanced distributed learning; sharable content object reference models and application program interfaces enhanced and extended the e-learning frameworks to a greater horizon. Present technology also ensures transformations of e-learning information without any geographical barriers. These educational and student or user data combined together forms big data architecture under e-learning environment and mining these big data for various requirements or knowledge discoveries needs innovative approaches. This paper identifies and evaluates various e-learning models and associated education technology paradigm. The research further explores and proposes a new framework for big data integrations. The paper also discusses the scope of future research on data mining and role of big data in e-learning environment.

Keywords: E-learning, Big Data Framework, Augmented learning, Adaptive learning

I. INTRODUCTION

Technology innovations and research in education technology contributed enhancement of e-learning framework to a newer dimensions. Research also showcased, how new methods of teaching and learning including student performance evaluations can be incorporated with in the defined e-learning framework[1]. Advancement in Internet technology and information and communication technology ensured storage, scalability and sharing of education information irrespective of geographical barriers.

Basic e-learning framework consists of user centric features, which are embedded over the e-content repositories [2]. Information from the e-content can be viewed or exchanged between user and the repositories using various user interfaces.

Education technology support and e-learning software tools and techniques further enhances the speedy available of these information as per selections and requirements [3]. Storage and retrieval facilitates storage large volume of e-learning information and selective or conditional retrieval as per user choice or requirements. Advancement in internet technology, wireless technology, server technology, data transmission and network technologies ensures proper sharing and transmission of e-learning information across the users irrespective of geographical barriers. Incorporating authentication with in the e-learning framework helps retrieval and viewing, transmission of information among authorized users only[4].

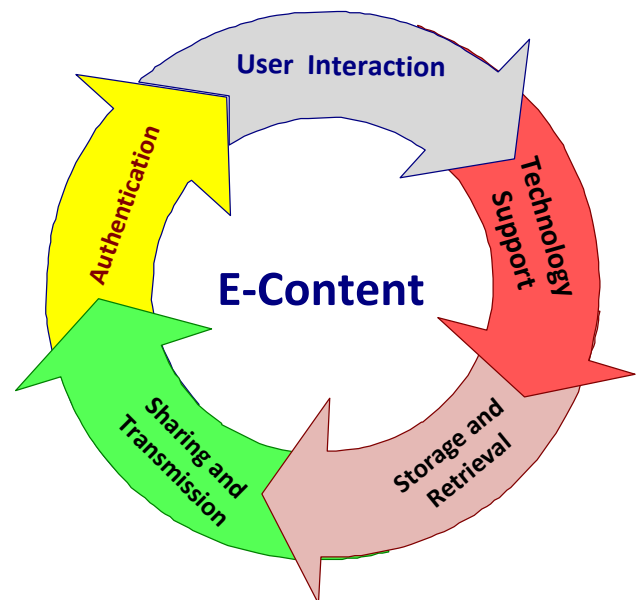


Fig. 1. Basic e-learning framework

II. STUDY OF EXISTING SYSTEM

E-learning systems are further enhanced using various education technology platforms and learning management

systems. Learning Management System integration over e-learning framework [5] is showcased in Figure 2.

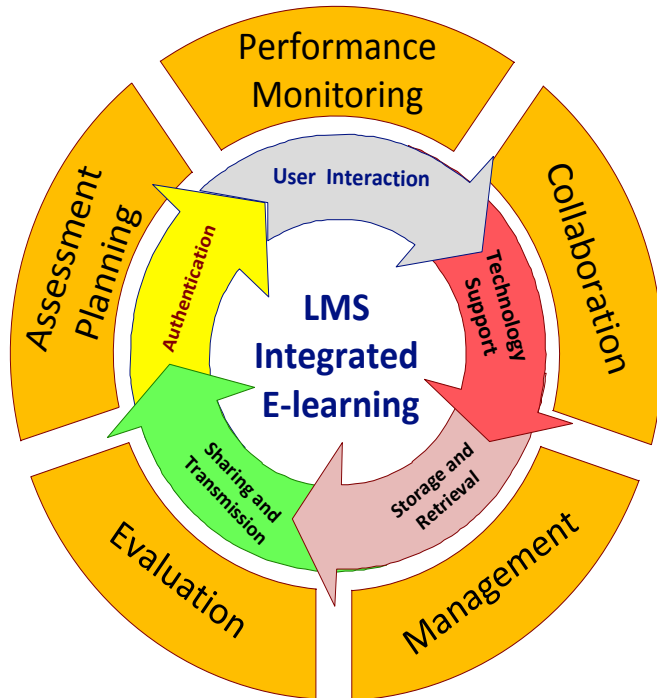


Fig. 2. Integrated Learning management System

Learning management system facilitates various features for managing e-learning resources along with students or user information. Few of the basic functions like course management, content management and course planning are efficiently handled by learning management systems. One of the complex process can be managed by learning management system is student performance monitoring. Course administration, feedback through dash boards and messages, automating various learning process, uploading and downloading facility of contents, replica of virtual class room along with multimedia audio, video, text, graphics, animations integrations are possible in new learning management systems [6].

Conduction of online tests, configuring for multiple attempts, test administrations, custom content development and integrations, publishing of results and reports, post exam reviews, learning group administrations are some of the key features of present day learning management systems. The system facilitates tracking of student performances during evaluation process. Student results can also be stored for future references [7].

Integration of sharable content object reference model (SCORM) with in learning management system facilitates student performance evaluation as well. SCORM is a set of standards defines how communication can be established between sharable contents and learning management systems as shown in figure 3. Introduction of SCORM further

enhanced learning management systems for better e-learning across various frameworks and online learning platforms [8]. Due to the very fact that now e-contents can be developed only once and can be shared across various systems.

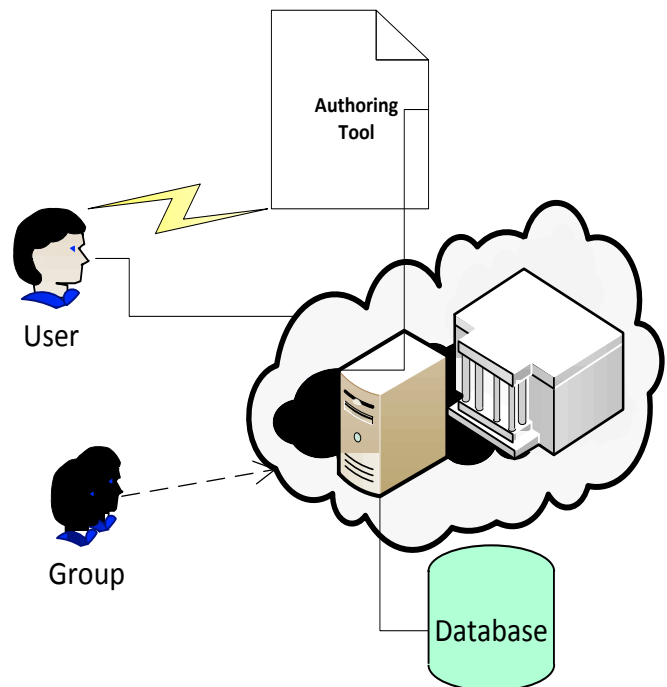


Fig. 3. SCORM integration in LMS

Basic authoring tools are necessary and must be incorporated with in the systems for content development for learning management systems. These authoring tools also facilitate conversion of contents with SCORM features for collaborative learning [9].

III. PROPOSED SYSTEM FOR EDUCATION DATA MINING

Data mining of education data from e-learning resources helps to extract useful patterns and their associated relationship. Large volume of e-learning data is normally unorganized and consists of varied dataset, when we fetch from e-learning framework for any analysis [10]. Systematic integration of data mining ensures re-orientation of these educational data sets and helps to identify and detects the right data for utilizations. Hence the process of data mining of educational data begins with identification and detection of pre-process data. These data are processed and created as refined data set and further classified as various data groups. Hence raw data, which are retrieved from various e-learning systems, e-contents, repositories, portfolios, web based tutorials, on-line learning systems are converted, organized and classified as data groups.

The next phase of education data mining is clustering of organized data for identification of similar groups as shown in figure 4. . Purpose of clustering is to discover distinct e-learning user groups, e-learning content groups or e-learning

resource utilization groups. Further sub- classification or sub-clustering can be done as per requirements. This identification of groups help to analyze and develop knowledge to design new set of contents, resources or new requirements of users or learners.

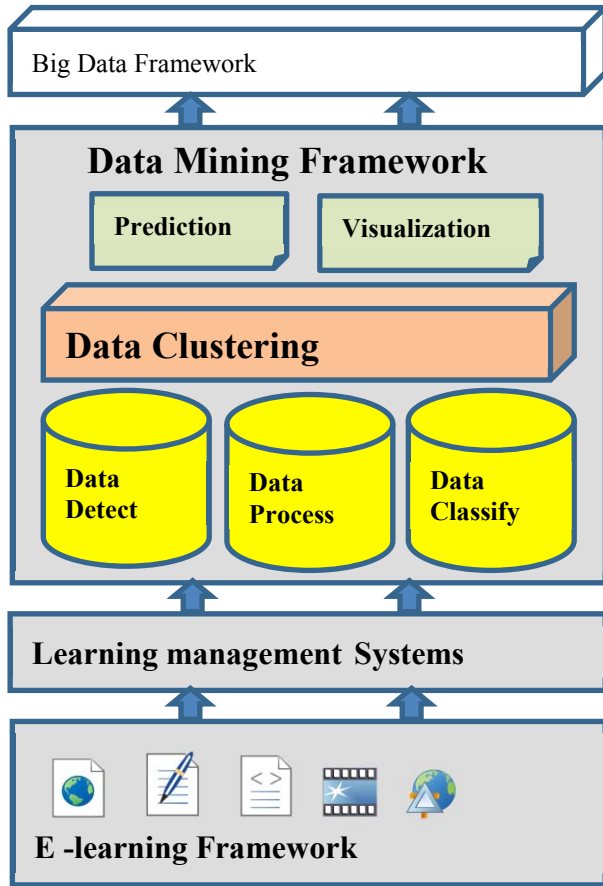


Fig. 4. Educational data mining framework

Recent developments in data clustering identified new ways of clustering for real time analysis. Multi dimension clustering helps to explore new features of grouping and multi-dimensional analysis of various results related to leanings as well as learners requirements are possible [11]. Different set of methods are available for data clustering. Few among them are decision rule, associated rule, decision tree method, conceptual cluster and taxonomy method.

Once the clustering is completed, the results are then passed further for pattern creation or pattern generations and visualizations phases. Predictive analysis can be done using the result of these phases [12]. Visualization of outcome using patterns, density graphs, predictive analysis, regression analysis, sequential pattern analysis, random pattern analysis, summarization can be done for visualization and analysis [13]. Here sequential pattern analysis helps to analyses frequently occurring events with reference to time or other events.

IV. BIG DATA INTEGRATION WITHIN THE E-LEARNING PROCESS

Big data integration after data mining is a part of knowledge discovery process. Internet technology influences, pervasive technology interventions on e-learning process resulted into real time transmission of education data constantly and continuously across the platforms. The end result of the big data due to continuous availability of data helps to discover various knowledge related to the specified domains as shown in figure 5.

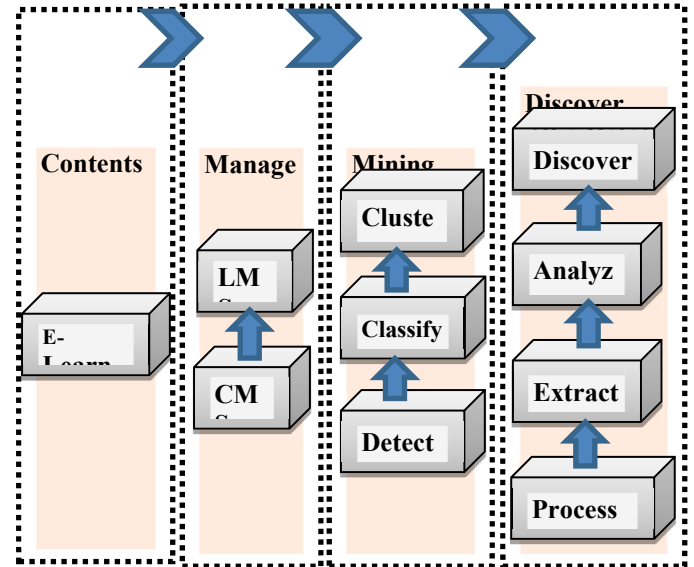


Fig. 5. Big data integration

Big data integration begins with processing of large volume of data passed from data mining phases. Once the data are properly processed and filtered feature extraction begins. Here various characteristics of data are identified and attributes like whether the information is numeric, data, text can be segregated group-wise for further analysis. Once the feature extraction process is completed, the information in the form of data is now ready for various predictive or statistical analysis. The outcome of analysis phase helps to discover various information pertains to different knowledge domains [14].

V. RESULT ANALYSIS

Integration of data mining capabilities and big data framework within e-learning process derives new set of capabilities, where along with content analysis, learning resource analysis, learners analysis also possible as shown in below table 1.

TABLE 1: e-Learning framework vs Big data features

Features	E-learning Portal	LMS	Data Mining & Big Data integration
E-content	√	√	√

Features	E-learning Portal	LMS	Data Mining & Big Data integration
display			
Interactive learning	√	√	√
Performance Monitoring	√	√	√
Couse Administration	X	√	√
Course Management	X	√	√
Performance Evaluation	X	√	√
Performance Tracking	X	√	√
Cluster Generation	X	X	√
Pattern Generation	X	X	√
Pattern Analysis	X	X	√
Predictive Analysis	X	X	√
Knowledge Discovery	X	X	√

Data mining and Big data framework integration exhibits superior outcomes because complex information analysis capabilities are possible, which are otherwise not possible in basic e-learning frame work or learning management system.

VI. CONCLUSION

The research outcome of this paper showcases the evaluation of basic e-learning framework. The paper further discusses the benefit of learning management systems with reference to e-learning paradigm. The paper also explores the possibility of education data mining features, embedding over the e-learning process and how the resultant outcome can be used in varieties of context. The research also discusses how the data mining concept helps to create clustering of data and how subsequent visualization and analysis can be done. The paper further gives a clear insight into the integration of big data framework within the e-learning framework and the resultant benefits out of big data integrations. The result clearly identifies how data mining and big data integration on e-learning process can derive the advance features such as pattern generation, pattern analysis, predictive analysis and knowledge discovery, which is needed to identify learning needs along with learners needs in future e-learning revolutions.

REFERENCES

- [1] Shima, B., Fatma, B., and Darwesh, G., " *Simulation-based performance evaluation of queuing for e-learning real time system*", Proceedings of International Conference on Education and e-Learning Innovations (ICEELI), Sousse, Tunisia, July 1-3 2012, pp. 1-6.
- [2] Kresimir, F., Bozic, N., and Hrvoje, J., "*Analysis of e-learning repository systems and frameworks with prepositions for improvements*", Proceedings of the 31st International Conference on Information Technology Interfaces (ITI, Dubrovnik, Croatia, 22-29 June 2009, pp. 487-492
- [3] Shanti D, *Learning new technology tools in pre-service teacher education: A model for instructional approach*, Proceedings ascilite 2011 Hobart, pp 352-353
- [4] Pierre G., Lionel K., *Management of the Transition to e-Learning in Higher Education Based on Competence Quotient*, Proceedings of International Conference on Communication, Management and Information Technology, Elsevier, 2015, PP 324-332, 2015
- [5] Georgoulis K., Skalkidis I., Guerreiro P., *A Framework for Adopting LMS to Introduce e-Learning in a Traditional Course*, Educational Technology and Society, 2008, 11 (2), pp 227-240
- [6] Nishantha G, Danh H. A., Davar P., *A Multimedia Integrated Framework for Learning Management Systems, eLearning - Theories, Design, Software and Applications*, Dr. Patrizia, Ghislandi (Ed.), 2012, pp 153-172..
- [7] Biswas P., Ghosh S.K., *A Novel Approach to Define Performance Metrics for Students' and Teachers' Evaluation*, The Electronic Journal of e-Learning, available online at www.ejel.org, Volume 5 Issue 2, pp 87 - 102.
- [8] Jonathan P., Nikolaus H., Andy J., Jason H., *The Next Generation of SCORM: Innovation for the Global Force, Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2012*, pp 1-9.
- [9] Chandrashekar R., Neha G., Mehul S., Aditya P., *A System for Managing SCORM-Compliant Learning Objects*, IEEE International Conference on Technology for Education, 2011.
- [10] Papamitsiou, Z., Economides, A., *Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence*. Educational Technology and Society, 17 (4), 2014, pp 49-64.
- [11] Pooja T., Anil M., Manisha., *Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue*, International Journal of Computer Applications (0975 – 8887) Volume 110 – No. 15, 2015, pp 60-68.
- [12] Amirah M. S., Wahidah H., Nur'aini A. R., *A Review on Predicting Student's Performance Using Data Mining Techniques*, Procedia Computer Science, Elsevier Volume 72, 2015, Pages 414-422
- [13] Brijesh K. B., Saurabh P., *Mining Educational Data to Analyze Students Performance*, International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011, pp 63-69.
- [14] Ahmed A., IHHazem E., IISamir M. A., Yehia E., *Handling Big Data in E-Learning*, International Journal of Advanced Research in Computer Science and Technology, Vol. 3, Issue 1, 2015, pp 47-5