

A COMPRESSIVE SURVEY ON E-LEARNING SYSTEMS USING EDUCATIONAL DATA MINING TECHNIQUES AND ITS SECURITY ISSUES

J S RADHIKA ^{*1}, RAGIPANI SOWMYA²

¹²Assistant Professor, Dept of CSE, Sri Indu College of Engineering and Technology, Hyderabad, Telangana, India

ABSTRACT:

E-learning represents a new context of education that persistently generates large measures of data that describe the coherence of interaction between learning and teaching. However, an abundance of available data is an exponentially developing problem, because unstructured and uncontrolled data might be obsolete or incorrect, and their sum congests the education system without giving any articulated knowledge. Information Mining is a process that uses measurable and mathematical procedures, techniques of counterfeit consciousness and machine learning to identify and extract useful data and subsequent knowledge from large databases to its members and is a suitable apparatus for tackling the problem of "data overload". In this paper, the utilization of the information mining methods on the classification of the common language sentences is presented. Recently, Educational Data Mining has become an emerging research field used to extract knowledge and discover patterns from E-Learning systems. This work is a survey of the specific utilization of information mining in learning management systems and a case contemplate with university database's-learning systems is a data superhighway where large-scale correspondence network is provided with a variety of interactive services are provided as text database, e-mail, sound, video etc. Cyberspace is the indefinite area broadly known as the Internet. Cyber security is the frame of conventions framed in one place for safeguarding this cyberspace. This paper explains how information mining undertakings like clustering can be applied to the information taken from an e-learning system and additionally deals with E-threats and E-risks associated with it. This paper will likewise center upon the pervasiveness of internal cyber-assault and additionally absence of proper IT policies and procedures in e-Learning systems. The performance of students on the online course in advanced electronics is taken for the investigation and results are achieved with WEKA instrument. introduced e-learning stage which has divided two vital part; the initial segment is learning information must be secured, for securing information we have used file encryption and decryption technique, and second part comes use of information digging techniques and concepts for huge information storage.

Keywords: EDM, Classification, Clustering, WEKA, E- Threats and E-Risk, file encryption and decryption.

I. INTRODUCTION

Over a wide variety of fields, information are being collected and accumulated at a sensational pace. There is a urgent need for a new generation of computational theories and instruments to help people in extracting useful data (knowledge) from the quickly developing volumes of computerized information. These theories and instruments are subject to the merging field of knowledge discovery in databases (KDD).KDD refers to the overall process of discovering useful knowledge from information, and information mining refers to a specific step in this process. Information Mining is the utilization of specific calculations for extracting patterns from information. The widespread application areas of information mining include bio-informatics, e-commerce, extortion detection and now in the field of education too. The information mining in the field of educational

research is known as Educational Data Mining (EDM). EDM tries to simulate a student model which might be used for the improvement of students by predicting the future trends. The sharing of data, joint effort and interconnectivity are the core elements of an e-learning system. Since e-learning system is open, distributed and interconnected then, the security becomes a critical challenge in order to ensure that interested and authorized on-screen characters just have access to the correct data at the appropriate time [1].

Protection against information control, fraudulent user authentication and compromises in confidentiality are essential security issues in e-learning. E-risk involves the risk at the time of the electronic exchange, whereas threat means an anticipated danger [2]. In any case, in order to have a secure transmission of the data being exchanged over the internet, one needs the concept of Network Security, which needs to take punitive activity to Ease of Use protect from different types of attackers like-hackers, interested computer neophytes, deceitful vendors or disenchanted employees of an association [3]. Basic threats to computers are viruses, network penetrations, theft and unauthorized alteration of information, eavesdropping, and nonavailability of servers and personal computers [4]. The challenge is the means by which to eliminate, manage and organize this enormous information. The genuine thing is to discover which information is essential, what to keep and what to dispose of [5].

This paper explains the methods of EDM through an example information taken from an e-learning stage and likewise deal with its security threats and risks. This paper is organized as: EDM and its imperative methods like classification, clustering. It likewise explains how WEKA apparatus is used to analyze the student information in online computerized electronics course taken from the university database.

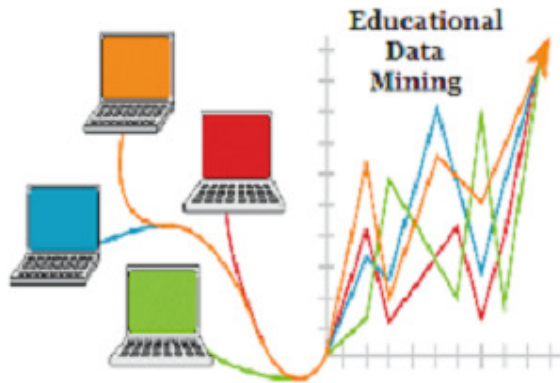
II. EDUCATIONAL DATA MINING

Educational information mining (EDM) describes a research field concerned with the use of information mining, machine learning and measurements to data generated from educational settings (e.g., universities and intelligent mentoring systems). At an abnormal state, the field seeks to develop and improve methods for exploring this information, which often has multiple levels of meaningful hierarchy, in order to discover new bits of knowledge about how people learn in the context of such settings.[1] In doing as such, EDM has contributed to theories of learning investigated by researchers in educational brain science and the learning sciences.[2] The field is closely tied to that of learning examination, and the two have been compared and contrasted.[3]

Educational information mining refers to techniques, devices, and research designed for naturally extracting meaning from large repositories of information generated by or related to people's learning activities in educational settings. Quite often, this information is extensive, fine-grained, and precise. For example, several learning management systems (LMSs) track data, for example, when each student accessed each learning object, how often they accessed it, and how long the learning object was displayed on the user's computer screen. As another example, intelligent coaching systems record information every time a learner presents an answer for a problem; they may collect the time of the accommodation, whether or not the arrangement matches the expected arrangement, the measure of time that has passed since the last accommodation, the order in which arrangement components were entered into the interface, etc. The precision of this information is to such an extent that even a genuinely short session with a computer-based learning environment (e.g., 30 minutes) may produce a large measure of process information for examination.

In other cases, the information is less fine-grained. For example, a student's university transcript may contain a temporally ordered rundown of courses taken by the student, the grade that the student earned in each course, and when the student selected or changed his or her academic major. EDM leverages the two types of information to discover meaningful data about different types of learners and how they learn, the structure of space knowledge, and the effect of instructional strategies embedded inside different learning environments. These analyses provide new data that would be hard to discern by taking a gander at the crude information. For

example, breaking down information from a LMS may reveal a relationship between the learning objects that a student accessed amid the course and their last course grade. So also, examining student transcript information may reveal a relationship between a student's grade in a specific course and their decision to change their academic major. Such data provides knowledge into the design of learning environments, which permits students, teachers, school heads, and educational policymakers to make informed decisions about how to interact with, provide, and manage educational resources.



- Content delivery through computer utilizing WWW
- It focuses on the broadest view of learning and learning arrangements.

The significant challenge in E-learning systems is the collection of unstructured data that chokes the educational system without giving any articulate knowledge to its performing artist 's information mining was destined to tackle problems like this.EDM is an emerging discipline, concerned with developing methods for exploring the unique types of information that come from educational settings, and utilizing those methods to better understand students and the settings which they learn in. EDM develops methods and applies techniques from measurements, machine learning and information mining to analyze information collected amid teaching and learning. Information Collected from online learning systems can be aggregated over a large number of students and can contain numerous variables that information digging calculations can explore for model building.

Objectives of EDM:

- Predicting student's future learning behavior by creating student models that incorporate such detailed data as students' knowledge, inspiration, metacognition, and attitudes.
- Discovering or enhancing space models that characterize the content to be learned and ideal instructional sequences.
- Studying the effects of different sorts of pedagogical help that can be provided by learning software and Advancing scientific knowledge about learning and learners through building computational models that incorporate models of the student, the space, and the software's pedagogy.

DATA MINING TECHNIQUES IN EDM

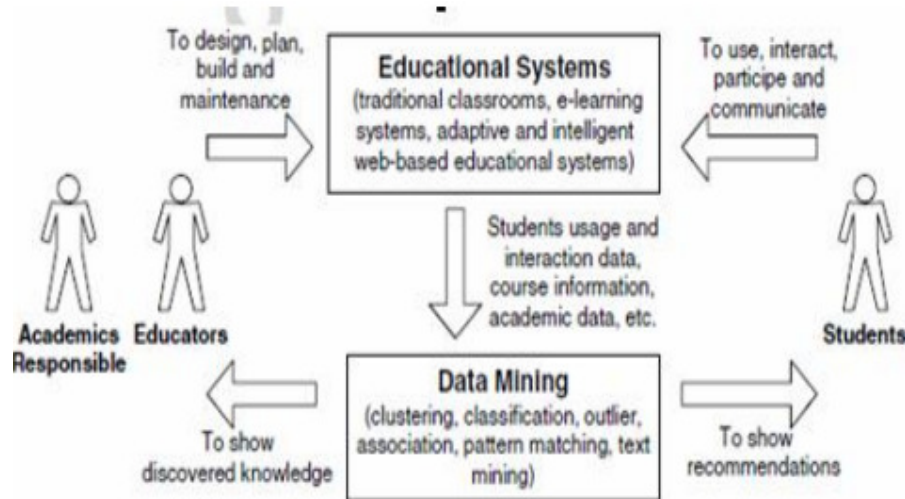


Fig.2. The cycle of Applying Data Mining in Educational systems

The utilization of information mining in e-learning systems: can help to resolve different problems utilizing different information mining techniques.

Information mining can be used to resolve classification problems in e-learning. Just a few information mining techniques can be applied to learning to resolve classification problem. The techniques are fluffy rationale methods; manufactured neural networks and evolutionary calculation, charts, and trees; affiliation rules; multi-agent and systems. Furthermore, use of information mining to resolve clustering problems in e-learning; includes counterfeit neural network and clustering.

Other information mining techniques that can be used in e-learning prediction techniques, perception techniques, and case-based reasoning. Information mining application in e-learning is an iterative cycle. The mined knowledge ought to enter the circle of the system and guide, facilitate and enhance learning as a whole, transforming information into knowledge as well as filtering mined knowledge for decision making. - In this work, we have indicated how useful the use of information mining techniques in course management systems can be for online educators. In spite of the fact that we have demonstrated these techniques separately, they can likewise be applied together in order to get interesting data in a more efficient and faster way. To start with, teachers can use perception techniques to acquire a general view of the student's usage information. And for example, on the off chance that they discover something strange or irregular in the plots, then they can acquire more detailed data about these events by viewing factual values. Or on the other hand, on the off chance that they locate some comparative gatherings of students in diagrams, then they can apply clustering techniques in order to acquire the exact gathering's students can be divided into. And these gatherings can likewise be used to create a classifier in order to arrange students. The classifier indicates what the primary characteristics of the students in each gathering are, and it enables new online students to be classified. At long last, the teachers can apply affiliation rule mining to discover if there is any relationship between these characteristics and other attributes. These rules can help to group students as well as to detect the sources of any incomprehensible values obtained by the students.

Processes of Data Mining in E-Learning

Cristóbal Romero et. al., [5] gave a detailed description of The use of information mining in e-learning systems is an iterative cycle in which the mined knowledge ought to enter the circle of the system and guide, facilitate and enhance learning as a whole, transforming information into knowledge as well as filter mined knowledge for decision making. The utilization of information mining in educational systems has specific requirements, mostly the need to take into account learners' specific behavior, including pedagogical aspects. The use of information mining in E-learning systems can be described as an iterative cycle where information mining applications contribute to enhancing learning, and additionally utilizing mined knowledge for decision making.

The e-learning information mining process comprises of the same four steps in the general information mining process (see Figure 3) as takes after:

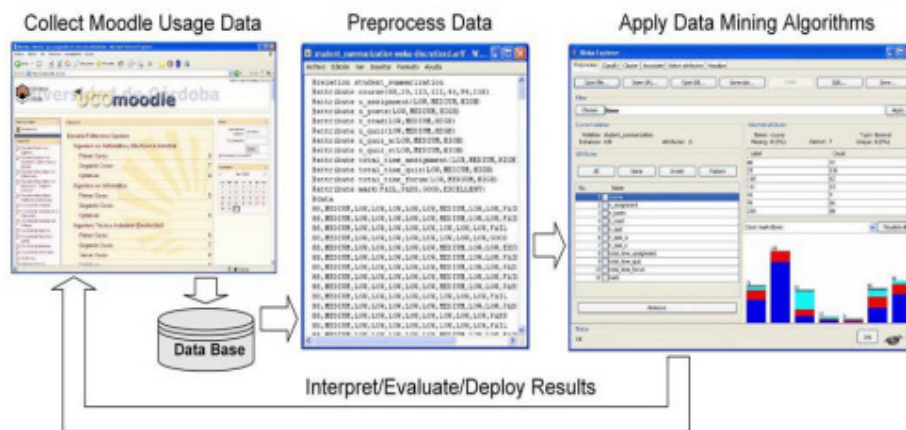


Fig.3. Mining Moodle data

1. Collect information. The LMS system is used by students and the usage and interaction data is stored in the database. In this work, we will use the students' usage information of the Moodle system. Interaction data is stored in the database of the LMS.
2. Preprocess the information. The information is cleaned and transformed into an appropriate organization to be mined. In order to preprocess the Moodle information, we can use a database executive apparatus or some specific preprocessing instrument. The information is transformed into an appropriate organization.
3. Apply information mining. The information mining calculations are applied to fabricate and execute the model that discovers and summarizes the knowledge of interest for the user (teacher, student, director, etc.). In order to do as such, we can use a general or a specific information mining instrument, and we can use a commercial or free information mining apparatus. The information mining calculations are applied to create and execute the model that discovers the knowledge and patterns of interest. In order to achieve this objective, an information mining device can be used.
4. Interpret, evaluate and deploy the results. The results or model obtained are interpreted and used by the teacher for further activities. The teacher can use the data discovered to make decisions about the students and the Moodle activities of the course in order to improve the students' learning. The model obtained is interpreted and used by the educator for further analyses. The educator can use the data discovered to make decisions about e-learning system and process.

CYBER SECURITY THREATS AND RISKS IN E LEARNING SYSTEMS

Security of Network these days is becoming a very basic facet when the degree of information to be shared is huge and complex. So the imperative issue today is the data secrecy for the secure information correspondence over the network; protecting information from being disclosed to illegitimate users while legitimate users are sharing the information. Cryptography in Networks helps for the same as it enhances the information confidentiality and protection by making data incoherent. Hence, the information or the data remains protected from the illegitimate users.

Information Encryption and decryption are the two most critical techniques used today to enhance the security of the information. Cryptography is a craft of converting the information on the applications into incoherent or scrambled or in a unintelligible configuration. It related to the investigation of mathematical calculations related to aspects of data security, for example, confidentiality, information integrity, and information authentication [6]. Symmetric Key Encryption and Asymmetric Key Encryption are other two essential encryption types [7].

Building trust and encouraging engagement among users of e-learning systems is essential and becoming vital because there are opportunities for both synchronous and offbeat interactions with the system. Synchronous learning happens in real-time, with all members interacting at the same time, while nonconcurrent learning is self-paced and enables members to engage in the exchange of ideas or data without the dependency of other participants' involvement at the same time [1].

A. E-threats-A threat in e-learning may cause e-risk. There are significant threats, for example, - Confidentiality infringement, Integrity Violation, Denial of service[1,4], Illegitimate Use [6,7], Malicious Program, Repudiation, Masquerade, Traffic Analysis, Brute-force assault, [1,4], Guessing passwords, seizing, Sniffers, Social Engineering, Spoofing [4]. Every one of these threats are to be dealt separately by halting the penetration and loopholes.

Due to above-mentioned threats following risks may happen. Before even beginning with any e-learning system risks ought to be given highest need.

B. E-risks-Institution ought to perform a cybersecurity risk assessment. There are numerous e-risks like creator's risk, Teacher 's risk, Manager 's risk, System developer's risk, Student, risk.

I. Writer 's risk-A writer is someone who is composing and developing the content in an e-learning system. Presently the real risk at creator's end is that there content and notes may get altered or processed without their knowledge. There content may be used in an unauthorized manner, modified and reused in a different context in e-learning.

ii. Teacher 's risk-teacher in an e-learning system is someone who delivers a lecture and uses material by different author's. The risk with teachers is the way to deliver a lecture, sending notes and assignments, accepting and stamping answer sheets, preparing and dispersing mark sheets.

iii. Manager 's risk-There are three types of risk at managerial level. Right off the bat, inelegant people dissimulating as students and composing the paper instead of unique students. Secondly, keeping up passwords of all servers and routers, recording of day by day movement and taking care of other network devices. Thirdly, to control approval i.e. access strategies.

iv. System Developer 's risk-there is two types of risk involved as developers end. Right off the bat, giving quality hardware components like high ended servers for designing, developing, and delivering E-learning items. Secondly, putting away passwords in clear text in the application code so that isn't accessed by the intelligent learner.

v. Student 's risk-receiving unaltered question paper, ought to be aware of phishing, and store login data in order to prevent unauthorized access.

E-learning Security techniques

E-learning has been divided into two sections: a. Information Security b. user flexibility. E-learning has a huge database which carries heaps of student records, course records, course materials and so on.

In this system user security provided by the administrator, administrator himself authorize the candidate to enter into the system. Course material likewise has been secured by utilizing file encryption and decryption technique with the goal that nobody can access material outside the stage. Let we see how cryptography in learning stage:

Cryptography is processed in which conventional text (plaintext) scrambling into ciphertext (this process recognized as encryption) then back again to plain text (this process is known as decryption). In proposed system file encryption and decryption technique is used with the goal that course material either pdf or video instructional exercise are can't be interrupted by the obscure user likewise materials can be used confidentially.

Open key/two-key/asymmetric cryptography involves the use of two keys:

- an open key, which might be known by anyone, and can be used to encrypt messages, and verify signatures
- a private-key, known just to the recipient, used to decrypt messages, and sign (create) signatures

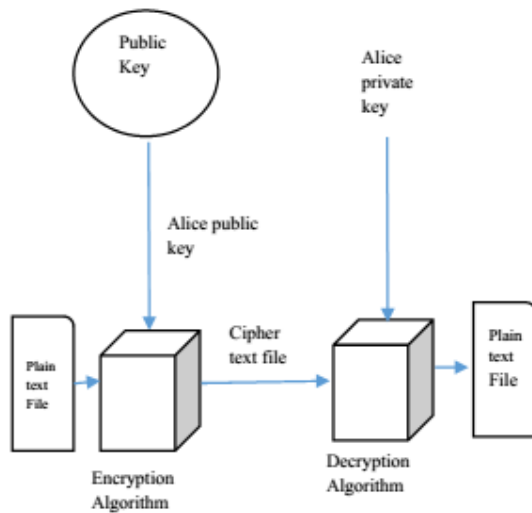


Figure 4. File encryption and Decryption

In E-learning stage secure file transmission done as takes after:

1. Just staff/Admin can transfer think about material to the E-learning stage.
2. Administrator selects file which will be uploaded.
3. File will be encrypted and stored to the system.
4. While learner wish to use that file then file will be retrieved.
5. Decrypted file provides to learner and moves to another area.

6. After learner sign out through the system decrypted file get cleared.

Encryption:

Let F be a file to be encrypted where the contents of file are taken into string S.

Select random number r, where $r < m$. Compute cipher text as: $c = g^{s^e \bmod n} \times r^m \bmod m^2$.

Decryption

Compute unique message:

$$S = (((c^3 \bmod m^2 - 1)/m) \times \mu \bmod m) \bmod n.$$

Above process helps to keep consider material confidential. For example if any third person entered to the system or either registered candidate downloads the any video instructional exercise then he unable to watch it because of encryption. So it make mandatory to learn online through the e-learning system. With the goal that it becomes easy to observe student

every day attendance or student behavior to the system. And additionally the system provides course length so learner includes complete learning inside term after that learner unable to enter to the system. On the off chance that learner required to enter to the system again then again he should register with new confirmation id to the system. Because of that system, learner and in addition information additionally has been secured in proposed E-learning System. Presently we turn towards another piece of this system that is user flexibility. E-learning is advanced stage so it provides whenever anywhere learning. Internet and computer based learning gives feel as single learner to each individual learner. These things provide learner happiest, easiest learning then additionally for more convenience information concepts has been involved in this E-learning system. We can provide scientific categorization of e-learning problems to which information mining techniques have been applied like: students classification based on their course, e-learning system routes and interaction improvements etc.

Information mining techniques, for example, Association rules, inter session and intra session were applied to extract useful patterns that may help educators, educational managers to evaluate and interpret online course action. The likelihood of following user behavior in e-learning stage makes possible the mining of the resulting information bases. The information mining methods applied to evaluate the learning material in an e-learning course.

E-Learning course offerings are presently some more, and numerous new e-learning stages and systems are developed or implemented. These stage generate an exponentially increasing measure of information, and a lot of this data can possibly become knowledge to improve all instances of e-learning. Information mining process should enable the extraction of this knowledge.

The information mining has been involved in ELearning has huge reason that is the wide capacity of huge measure of information and imminent need for transforming information into useful data or knowledge. The use of information mining in educational systems has specific requirements, principally the need to take into account learners' specific behavior, including pedagogical aspects. The utilization of information mining in E-learning systems can be described as an iterative cycle where information mining applications contribute in enhancing learning, and likewise utilizing mined knowledge for decision making. In studies on how Data Mining techniques could successfully be incorporated to e-learning environments and how they could improve the learning errands were carried out. Information clustering was suggested as a means to promote gather based collaborative learning and to provide Incremental student analysis.

CONCLUSIONS

In this work, we have demonstrated how useful Data Mining Techniques can be if applied to Learning Management Systems. They help in dissecting/predicting student behavior pattern and update E-learning systems as needs be to benefit students, educator, system developer etc. With E-learning systems spreading its wings in the education sector to reach greatest people the security of such systems ought to be set to the best need. Further, the level of security could be secured by utilizing different cryptographic techniques, DRM, biometric authentications, access control utilizing firewalls etc.

REFERENCES

1. Bandra, F. Ioras, K. Maher, "Cyber Security Concerns in ELearning Education," International Conference of Education, Research and Innovation Seville, Spain. Proceedings of ICERI2014 Conference, Nov 17-19, 2014, Seville, Spain, ISBN: 978-84-617-2484-0, ISSN: 2340-1095, Published: IATED, pp. 0728-0734. [Online], Available: http://ecesm.net/sites/default/files/ICERI_2014.pdf
2. N. Barik, Dr. S. Karforma, "Risks and Remedies in E-learning System," International Journal of Network Security & Its Applications (IJNSA), vol. 4, No. 1, pp. 51-59, Jan 2012. DOI: 10.5121/ijnsa.2012.4105. [Online], Available:<http://arxiv.org/ftp/arxiv/papers/1205/1205.2711.pdf>
3. Francesco Maiorana, Angelo Mongioj, and Marco Vaccalluzzo(2012). Proceedings of the World Congress on Engineering 2012 Vol I WCE 2012, July 4 - 6, 2012, London, U.K. [2]
4. Aakanksha Chopra, "Security Issues of Firewall," International Journal of P2P Network Trends and Technology (IJPTT)- vol. 22, no. 1, Jan 2016, ISSN: 2249-2615, pp. 4-9. [Online], Available: <http://www.ijettjournal.org>
5. Zeng-gang, X., &Xue-min, Z. (2010). Research and Design on distributed Firewall based on LAN. Computer and Automation Engineering (ICCAE), 2010, E-ISBN: 978-1-4244-5586-7, Print ISBN: 978-1-4244-5585-0, INSPEC Accession Number: 11259785, DOI: 10.1109/ICCAE.2010.5451596,pp: 517-520, Publisher: IEEE, Singapore.
6. Suman Madan, Aakanksha Chopra, "Data Driven Analytics: Prospects and Challenges," Advances in Computer Science and information Technology (ACSIT), Print ISSN: 2393-9907, Online ISSN: 2393-9915, vol. 2, no. 4, April-June, 2015.
7. Aakanksha Chopra, Manish Singh, "Key Exchange Algorithms in Cryptography: A roadmap to Data Security," International Conference organized by JIMS
8. Aakanksha Chopra, "Comparative Analysis of Key Exchange Algorithms In Cryptography and its Implementation," Journals of Innovations- IMS Noida, vol. VIII, issue-2, Dec 2013, Print ISSN: 0974-7141, Online ISSN: 0976-1713.
9. Cristóbal Romero *, Sebastián Ventura, Enrique García"Data mining in course management systems: Model case study and tutorial, Department of Computer Sciences and Numerical Analysis, University of Córdoba, 14071Córdoba, Spain Nov17-19, 2014, Seville, Spain, [Online], Available:[ww.elsevier.com/locate/compedu](http://www.elsevier.com/locate/compedu)
10. Ankita Chopra," Quantitative Analysis of Dairy Product packaging with the application of data mining techniques," International Journal of Computer science and Information Technology, ISSN: 0975-9646, vol 7, no 2, Mar-Apr, 2016.
11. [Francesco Maiorana, Angelo Mongioj, and Marco Vaccalluzzo(2012). Proceedings of the World Congress on Engineering 2012 Vol I WCE 2012, July 4 - 6, 2012, London, U.K. [2]
12. Dr.P.Nagarajan1. Dr.G.Wiselin Jiji(2010): Online Educational System (e – learning), International Journal of u-and e- Service, Science and Technology Vol. 3, No. 4, December, 2010
13. Devaji Mahanta, Majidul Ahmed published paper "E- learning Objectives , Methodologies, Tools and its Limitations" , ISSN: 2278-3075,Volume-2, Issue-1, December 2012
14. Felix Castro, Alfredo Veligo, Angelo Nebot,Francisco Mugica "Applying Data Mining Techniques to e-earning problems" sci2u.ugr.es/keel/pdf/specific/capitol/ ApplyingDataMiningTechniques.pdf
15. Baker M.: The Roles of Models in Artificial Intelligence and Education Research: A Prospective View, International Journal of Artificial Intelligence inEducation 11(2000) 122-143.

16. Chu K., Chang M., Hsia, and Y.: Designing a Course Recommendation System On Web Based On The Students Course Selection Records. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications (2003) 14-21.
17. Margo H.: Data Mining in the E-Learning Domain. Computers & Education 42(3) (2004) 267-287.
18. Tavagrian D., Laypold M., Nolting K., Roser M., (2014). Is e-Learning the solution for individual learning? Journal of e-Learning, 2004.
19. Ajayi L.A. (2008), towards effective use of information and communication technology for teaching in Nigerian colleges of education, Asian J. Inf. Technol 7(5):210-214.