A Proposed Data Mining Framework for Higher Education System

Ayman E. Khedr
Department of Information Systems
Faculty of Computers and Information, Helwan University

Ahmed I. El Seddawy
Department of Business Information System
Arab Academy for Science and Technology

ABSTRACT

Educational data mining is concerned with the development methods for exploring the unique types of data that come from the educational context. Furthermore, educational data mining is an emerging discipline that concerned with the developing methods for exploring the unique types of data that come from the educational context. This study focuses on the way of applying data mining techniques for higher education system by using the most common techniques on most common application called Moodle system in education system. There are an increasing numbers of researches that interest in using data mining in education system. The proposed system for Higher Educational Data Mining System (HEDMS) is concerned with the developing methods that discover useful knowledge from data that extracted from educational system. The data collated form historical and usage data reside in the databases of educational institutes. The proposed system helps to get sufficient results which consist of several steps in our case study starting with collected data, pre-processing, applying data mining techniques and visualization results. We collected students' data from Moodle database.

Keywords

Moodle, VLE, LMS, MLE, EDM, DSS, DM, MIS, Clustering, Classification, Association Rule, K-Mean, Olapand Visualization.

1. INTRODUCTION

Basic concept of managing the knowledge is not a new issue however it is still developing its theoretical home. Data Mining has always been conducted in one way or another, e.g. apprenticeships, colleagues chatting or a parent handing over her/his business to the offspring. E-learning is one of the tools used in the knowledge management to share the knowledge among groups. E-learning is an approach to facilitate and enhance learning through ICT [1].

The proposed system 'HEDMS' will support the instructors in their making decision in the right time. This study aims at investigating the adoption process of decision making for instructors to support their students in a way of delivering any course related to any scientific department before any exams that could be classified and clustering the students based on usage of Moodle as in our case study.

The rest of paper is divided into following sections; section two is a background which is divided into two subsections, section one is about E-Learning, section two is about DM. Section three focuses on the literature review and significant prior research. Section four presents the proposed Higher Education Data Mining System 'HEDMS. Section five illustrates the finding and analysis results divided also into two subsections started with experiments and data pre-

processing. Finally, section six present conclusion and future works.

2. BACKGROUND AND RELATED WORK

During the past decades, e-learning has become the most important part of the global learning process and step by step revolutionized educational environments. There are many synonyms for the term "e-learning" that are used in different researches. The study on hand will adapt the terms computer-based learning technology-enhanced learning or "networked learning" [1]. Consequently, the term "e-learning" will be used that could be understood as combining the other ones mentioned. E-Learning is a very broad term for internet-based learning in general. Distance education, online learning and E-Learning all of these terms are becoming synonymous with the latest approach to providing high quality educational offerings. E-Learning places the greater accent in the interaction and communication between instructors and their students via internet-channels, videoconferencing, in asynchronous email or bulletin board sessions or synchronous such as chat room, whiteboard, application sharing sessions [5]. Moreover, the term "E-Learning" has created publicity over the last couple of years and implies commercialization aspects. The Internet has allowed colleges, corporate universities and for-profit businesses to begin offering degrees and executive education via the web. This approach thanks to that the companies offering computer based training established so-called E-Learning solutions. One of the largest companies in this business field, CBTS renamed itself introducing the trademark "The E-Learning Company" [5].

Self-organized learning, focus on meta cognitive learning strategies and preparing lifelong learning skills, emphasizing on process-oriented learning to focus on "learning to learn" instead of product oriented learning, shift from training to self-responsible learning. Student and team-oriented methods and collaborative learning based on constructive learning theories are building a community of learners, experts, facilitators, coaches, high flexibility personalized and individualized learning according to different learning types and personal preferences [5].
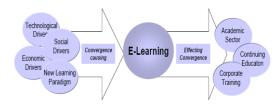

Fig. 1: E-Learning Convergence

E-Learning represents the convergence of several factors. Conversely, the growth of E-Learning and distance education has converging effects on the educational market as depicted in fig.1. The once almost visible line between the academic and corporate sector is disappearing. Companies and universities become both customers and suppliers for the new E-Learning business models [5].

2.1 E-Learning

E-learning therefore is considered as a best approach to facilitate and enhance learning through and based on both computer and communications technology. Such devices can include personal computers, CD-ROMs, Digital Television, PDAs and mobile phones. Communications technology enables the use of the Internet, email, discussion forums, collaborative software and team learning systems [1].

The use of web-based education systems has grown exponentially in the last few years, encourage by the fact that neither students nor teachers are bound to a specific location and that resulted of computer-based education which is virtually independent of any specific hardware platforms [1]. Specifically, collaborative communication tools are becoming a widely used in educational contexts. So, the Virtual Learning Environments 'VLE' are currently installed more and more by universities, community colleges, schools, businesses, and even individual instructors in order to add web technology to their courses and to supplement traditional face-to-face courses [2]. Such e-learning systems are also known as a Learning Management System 'LMS', Course Management System 'CMS', Learning Content Management System 'LCMS', Managed Learning Environment 'MLE', Learning Support System 'LSS' or Learning Platform 'LP'. These systems can offer a great variety of channels and workspaces to facilitate information sharing and communication between participants in a course that support educators to distribute information to students, produce content material, prepare assignments and tests, engage in discussions, manage distance classes and enable collaborative learning with forums, chats, file storage areas, news services, etc. [1][2].

Nowadays, one of the most commonly used is Moodle, Modular Object Oriented Developmental Learning Environment, which is a free learning management system that enables the creation of powerful, flexible and engaging online courses and experiences [3]. These e-learning systems accumulate a vast amount of information which is very valuable for analyzing students' behavior and could create a gold mine of educational data [3]. Learning management systems accumulate a great deal of log data about students' activities. These systems can record whatever student activities are involved such as reading, writing, taking tests, performing various tasks, and even communicating with others.

The traditional development of e-learning courses is a difficult activity. The system developers usually use the course teacher to choose the contents that will be studied and decide on the structure of the contents that determine the most appropriate content elements for each type of potential student in the course. Due to the complexity of these decisions, a one-shot design is hardly feasible, even when it is carefully done. As an alternative, it will be necessary in most cases to evaluate and possibly modify the course contents, structure and navigation based on students' usage information, preferably even following a continuous empirical evaluation approach [2].

Educational Data Mining 'EDM' is an emerging discipline, concerned with developing the appropriate methods for exploring the specific types of data that come from educational settings. Those methods are used to better understand students and settings what the students learn in. A key area of EDM is mining computer logs of student's performance. Another key area is mining enrollment data. Key uses of EDM include predicting student's performance and studying learning in order to recommend the improvement ways on the current educational practice. EDM can be considered one of the learning sciences, as well as an area of data mining. A related field is learning analytics [1].
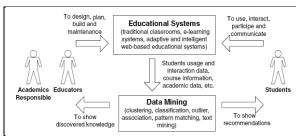

Fig2: Educational Data Mining [17]

2.2 Data Mining Techniques (DM)

Data mining is the process of analyzing data from different perspectives and summarizing it into useful information [13]. DM techniques are the result of a long process of research and product development [13]. The evolution of DM is shown in table 1.There is several processes for applying DM [6]:

1. Definition of the business objective and expected operational environment.

2. Data selection is required to identify meaningful sample of data.

3. Data transformation that involves data representation in an appropriate format for mining algorithm.

4. Selection and implementation of data mining algorithm depends on the mining objective.

5. Analysis of the discovered outcomes is needed to formulate business outcomes.

6. Representing valuable business outcomes.

Table 1. The evolution by DM [6]

Evolutionary Step	Business Question	Enabling Technologies	Product Providers
Data Collection (1960s)	"What was my total revenue in the last five years?"	Computers, tapes, disks	IBM, CDC
Data Access (1980s)	"What were unit sales in New England last March?"	Relational databases (RDBMS), Structured Query Language (SQL), ODBC	Oracle, Sybase, Informix, IBM, Microsoft

Data Warehousing & Decision Support (1990s)	"What were unit sales in New England last March? Drill down to Boston."	On-line analytic processing (OLAP), multidimensional databases, data warehouses	Pilot, Comshare, Arbor, Cognos, Microstrategy
Data Mining (Emerging Today)	"What's likely to happen to Boston unit sales next month? Why?"	Advanced algorithms, multiprocessor computers, massive databases	Pilot, Lockheed, IBM, SGI, numerous startups (nascent industry)

DM consists of five major elements; to extract, to transform, and to load transaction data onto the data warehouse system, to store and manage the data in a multidimensional database system, to provide data access to business analysts and information technology professionals, analyze the data by application software, and finally to present the data in a useful format, such as a graph or table [11].

DM techniques usually fall into two categories, predictive or descriptive. Predictive DM uses historical data to infer something about future events. Predictive mining tasks use data to build a model to make predictions on unknown future events. Descriptive DM aims at finding patterns in the data that provide some information about internal hidden relationships. Descriptive mining tasks characterize the general properties of the data and represent it in a meaningful way. Figure2 shows the classification of DM techniques.
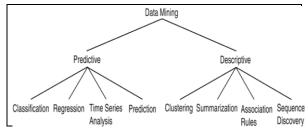

Fig 3: DM Techniques [5]

Association Rule is used to discover relationships between attribute sets for a given input pattern. Merceron [7] define sequence discovery as "a sequential technique is a given set of sequences find the complete set of frequent subsequences". Clustering is "the process of organizing objects into groups whose members are alike in some way" [7]. A cluster is therefore a collection of objects which are "similar" among them and are "dissimilar" to the objects belonging to another cluster. So, it deals with finding the internal structure in a collection of data.

Merceron [8] define that "Clustering involves identifying a finite set of categories or segments 'clusters' to describe the data according to a certain metric". Whereas, Kamber [9] has defined the clustering as "the technique that enables to find specific discriminative factors or attributes for the studied data. Each member of a cluster should be very similar to other members in its cluster and very dissimilar to other clusters. When a new data is introduced, it is classified into the most similar cluster".

Several researchers classified clustering algorithms differently. Some classifies clusters as mutually exclusive, hierarchical or overlapping. Others classifies cluster into hierarchal and partitioned. The most common classification is depicted in figure 5. Techniques for creating clusters include partitioning methods as in k-means algorithm, and hierarchical methods as in decision trees, and density-based methods.
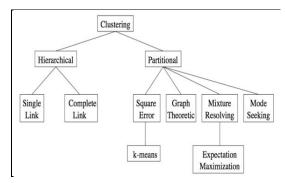

Fig 5: Clustering methods classifications [10], [14] and [15]

3. LITERATURE REVIEW AND SIGNIFICANT PRIOR RESEARCH

This section focuses on how data mining and how it could support e-learning processes and develop the education system via using several techniques and algorithms. This section appears as a survey for previous work in e-learning using data mining techniques. As stated in the introduction, the study aims at organizing the findings of the survey from different views that might correspond to the diverse academic researchers or professional backgrounds. In this section, the findings of research survey will present according to the data mining techniques such as classification, clustering, etc., algorithms and methods such as neural networks, genetic algorithms, decision trees, or fuzzy logic. In fact, most of the existing researches address the different solutions of classification and clustering. For this reason, specific subsections will be devoted to them. But first, let us try to find a place for data mining in the world of e-learning.

Data mining means searching for certain patterns within large sets of data, which creates a lot of possibilities for decision makers. By analyzing those patterns, better decisions can be made in order to improve learning and assessment process. The research interest in using DM in e-learning is constantly increasing. According to Mamcenko [25][26], to choose right DM techniques for analysis of students' examining data in order to explore students' behavioral characteristics, whereas having the exam electronic way and in accordance with the results to offer recommendations for a higher quality of exam arrangement and organization of examination [22]. Consistent with Sonali [23] to create and share our various experiences of using data mining for education, especially to support reflection on teaching and learning, and to contribute to the emergence of stereotypical directions. Conferring to Sonali [23] develop a trust on DM techniques so that present education and business system may adopt this as a strategic management tool. A Decision tree approach is proposed in which may be taken as an important basis of selection of student during any course program. As stated by Mamcenko [25], he suggested a "CHi-squared Automatic Interaction Detector 'CHAID' based approach to analyze the performance of higher secondary students".

Earlier studies reveals that a few parameters like students school type, location, family background, educational organization, medium of teaching like Hindi or English are prime key factors to predict their performance in higher education [19]. According to Chen [12], he proposed a Bayesian Networks based approach for student data classification [18]. The research work implement binary data classification, 5 levels data classification and a regression based analysis for predictive performance calculation of classifications on algorithms and descriptive knowledge analysis for better quality management in educational organization. Chen [12] discussed a decision tree approach based on C4.5, id3 and Naïve Bayes. The database of learning management system includes much useful information which can be effectively used for the improvement of e-learning process [19]. Those teaching management systems accumulate a vast amount of information which is very valuable for analyzing the students' behavior and could create a gold mine of educational data [22]. Mamcenko [25] has presented a case study that uses data mining methods to identify behavior of failing students to warn students at risk before final exam.

According to Gartner [14], the comparison between different classification algorithms and check which algorithm is the optimal approach for classifying students' based on their final grade and other authors use neural networks for predicting student's marks. Descried association rules are very useful in educational data mining since they extract associations between educational items and present the results in an intuitive form to the teachers. The choice of DM method and its realizable algorithm depends on available data, set research goals and intended results [22]. According to Khedr [26] the database of learning management system includes much useful information which can be effectively used for the improvement of e-learning process. Emphasis that learning management systems accumulate a huge volume amount of data which is very valuable for analyzing the students' behavior and could create a gold mine from educational data [6][26]. Authors emphasize on the vast quantities of data that could be daily generated, it is very difficult to analyze this data manually and a very promising approach towards this analysis objective is the use of DM techniques using DM methods many kinds of knowledge can be discovered.

The discovered knowledge can be used to better understand students' behavior, to access student's learning style, to adapt course content according to student's knowledge and abilities, to assist instructors, and to improve learning and teaching process. Mamcenko [25] presents a case study that uses DM methods to identify behavior of failing students to warn students at risk before final exam [20].

Prior studies describe various algorithms of DM methods. The choice of DM method and its possible algorithm depends on available data, set research goals and intended results. In line with Mamcenko [25], association rules are very useful in educational DM since they extract associations between educational items and present the results in an intuitive form to the teachers. Highlight that association rules require less extensive expertise in DM than other methods [25]. There are several e-learning projects are used in DM techniques as shown in table 2.

Table 2. E-learning projects in DM Techniques [10], [25]

Project Name	DM tech	E-learning Topic	University	Link
LON-CAPA	*KNN, decision tree, K-M, EM, Statistical analysis and classifications segments, SW agents, Learning Association rules, prediction, Consistency queries(CQ)*	*Assessment system, feedback to e-learning actor and student behavior tracking*	Michigan state	www.lon-capa.org
ATutor			University of Canada	www.autor.ca
LExiKON			University of Nacional	www.lexikon.dfki.de
aLFanet			Spain	www.lexikon.org
AHA!			Eindhoven university	www.aha.win.tue.nl
WebCT			Webct	www.webct.org
Blackboard			Blackboard	www.blackbord.org

Some researchers have pointed out the close relation between the fields of Artificial Intelligence 'AI' and Machine Learning 'ML' -main sources of DM techniques and methods- and education processes., the author establishes the research opportunities in AI and education on the basis of three models of educational processes: models as scientific tool, are used as a means for understanding and forecasting some aspect of an educational situation [13].

Models as component or responding to some characteristic of the teaching or learning process are used as "a component of an educative arte fact and base models for design of the educational arte facts, assisting the design of computer tools by providing design methodologies and system components or by constraining the range of tools that might be available to learners" [15]. Studies on how DM techniques could successfully be incorporated to e-learning environments and how they could improve the learning tasks were carried out data clustering was suggested as a means to promote group-based collaborative learning and to provide incremental student diagnosis [15].

A review of the possibilities of the application of Web Mining, Web usage mining and clustering, techniques to meet some of the current challenges in distance education was presented. The proposed approach could improve the effectiveness and efficiency of distance education in two ways on the one hand, the discovery of aggregate and individual paths for students could help in the best organize the educator organization's courseware. On the other hand, virtual knowledge structure could be identified through web mining methods.

The discovery of association rules could make it possible for web-based distance tutors to identify knowledge patterns and reorganize the virtual course based on the patterns discovered [12].

4. HIGHER EDUCATIONAL DATA MINING SYSTEM 'HEDMS'

The proposed system called Higher Educational Data Mining System 'HEDMS' concerns with developing methods that discover knowledge from data and usage come from educational system. The objective of HEDMS is to build a DM system for academic instructor for education system.

HEDMS consists of several components; data gathering, preparing data to discover knowledge, data pre-processing, using data mining techniques in sequences steps start with classification data, clustering data especially using K-mean 'K-M' algorithm and enhanced K-mean 'E-K-M' algorithm to set which best result, then post processing and finally get result and visualize result to create best decision to take a good decision for instructor. HEDMS Shown in figure 5.There are proposed system will describe and mining e-learning data to apply the main DM techniques such as classification, clustering, association rule mining and

visualization of Moodle data. The e-learning DM process consists of four steps as follows shown in figure 6:

1. Collect data (The students' usage data of the Moodle system).

2. Preprocess the data (The data is cleaned and transformed).

3. Apply data mining (The data mining algorithms are applied to build and execute the model that discovers and summarizes the knowledge using data mining tool).
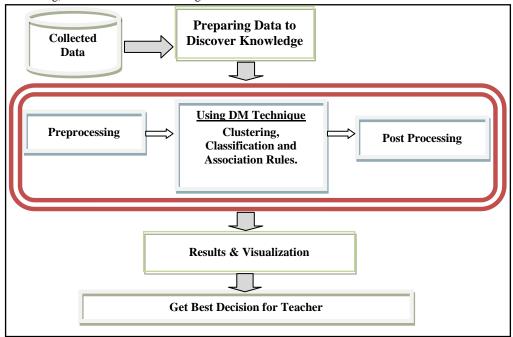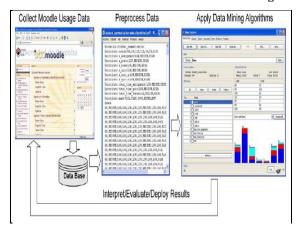

Fig 6: Proposed HEDMS

Fig 7: Mining Moodle Data [25]

Hardware for applying the HEDS is a personal computer configurations with this Processor 3.2, Hard Disk 160 gaga, Ram 2 G and Monitor 17 Inch.

Operating system is windows XP services pack 3. Several software tools have been used. The first is Microsoft Excel sheets 2007 and has been used for analysis and filtering data. Mat-lab version 6.5 has been used in data pre-processing and data classification. The last software is the WEKA which is a collection of java tools for DM written by staff at the University of Waikato, New Zealand [25].

5. FINDING AND ANALYSIS RESULTS

In this section, we present an evaluation of the proposed system educational in system using Moodle data. As data mining tasks needs pre-steps before applying the first step clustering algorithms, therefore, the following subsections will demonstrate the main steps in sections 5.1 and 5.2. In addition, section 5.3 will describe the finding of proposed system.

5.1 Experiments

Moodle is an open-source learning course management system to help educators create effective online learning communities. Moodle is an alternative to proprietary commercial online learning solutions, and is distributed free under open source licensing. Moodle has been installed at universities and institutions all over the world. Moodle has a large and diverse user community in over 75 languages in over 160 countries and more than 7,000 sites. An organization has complete access to the source code and can make changes if need be. Moodle can range from a single-teacher site to a 40,000 student university. Moodle's modular design makes it easy to create new courses, adding content that will engage learners [8].

Moodle is designed to support a style of learning called social constructionist pedagogy [18].This style of learning believes that students learn best when they interact with the learning material, construct new material for others, and interact with other students about the material. Moodle does not require the use of this style in the courses but this style is what it best supports. Moodle has a flexible array of module activities and resources to create five types of static course material (a text

page, a web page, a link to anything on the Web, a view into one of the course's directories and a label that displays any text or image), six types of interactive course material (assignments, choice, journal, lesson, quiz and survey) and five kinds of activities where students interact with each other (chat, forum, glossary, workshop and wiki)
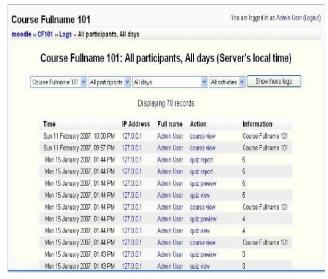

Fig.8: Moodle log report

HEDMS collected data from log file of database SQL2005. Moodle does not store logs as text files. Instead, it stores the logs in a relational database. So, data are stored in a single database. We have used MySQL because it is the world's most popular open source database (MySQL, 2007) to cane read and manage well. The Moodle database has about 145 interrelated tables, example of tables' uses in mining as shows in figure 9.

Name	Description
mdl_user	Information about all the users.
mdl_user_students	Information about all students.
mdl_log	Logs every user's action.
mdl_assignement	Information about each assignment.
mdl_assignment_submissions	Information about assignments submitted.
mdl_chat	Information about all chatrooms.
mdl_chat_users	Keeps track of which users are in which chatrooms.
mdl_choice	Information about all the choices.
mdl_glossary	Information about all glossaries.
mdl_survey	Information about all surveys.
mdl_wiki	Information about all wikies.
mdl_forum	Information about all forums.
mdl_forum_posts	Stores all posts to the forums.
mdl_forum_discussions	Stores all forums' discussions.
mdl_message	Stores all the current messages.
mdl_message_reads	Stores all the read messages.
mdl_quiz	Information about all quizzes.
mdl_quiz_attempts	Stores various attempts at a quiz.
mdl_quiz_grades	Stores the final quiz grade.

Fig 9: Example of tables' uses in Mining

5.2 Data Pre-processing

Data pre-processing allows the original data to be transformed into a suitable shape to be used by a particular data mining algorithm or framework. So, before applying a data mining algorithm, a number of general data pre-processing tasks can be addressed (data cleaning, user identification, session identification, path completion, transaction identification, data transformation and enrichment, data integration, data reduction)

Name	Description
course	Identification number of the course
n_assigment	Number of assignments handed in.
n_quiz	Number of quizzes taken.
n_quiz_a	Number of quizzes passed.
n_quiz_s	Number of quizzes failed.
n_messages	Number of messages sent to the chat.
n_messages_ap	Number of messages sent to the teacher.
n_posts	Number of messages sent to the forum.
n_read	Number or forum messages read
total_time_assignment	Total time spent on assignment.
total_time_quiz	Total time used in quizzes.
total_time_forum	Total time used in forum.
mark	Final mark the student obtained in the course.

Fig 10: example of attributes use for each student

Create summarization tables. It is necessary to create a new table (mdl_summarization) in the Moodle database that can summarize the information at the required level (e.g. student).

```
@relation student_summarization
@attribute course numeric
@attribute n_assignment numeric
- - -
@attribute total_time_forum numeric
@attribute mark numeric
@data
218,10,1,0,1,1,0,415,0,1417,6.0
218,0,0,0,2,2,0,0,641,1032,3.5
218,11,1,0,7,5,2,890,4667,5414,6.0
218,11,1,0,8,7,1,503,4735,4413,5.5
218,10,0,0,2,2,0,870,1237,100,2.5
218,10,0,0,6,5,1,332,1996,721,5.25
- - -
```

Fig 11: Example of student summarization

HEDMS will focus more on several columns such as, student ID and several actions or activates from students on course like (wiki, chat, messages, quizzes, forum, feedback, assignments and workshop). An activity is a general name for a group of features in a Moodle course. Usually an activity is something that a student will do that interacts with other students and or the teacher.

There are 14 different types of activities in the standard Moodle that can be found on the "add an activity" drop down menu, such as; Assignments: Enable teachers to grade and give comments on uploaded files and assignments created on and off line, Chat: Allows participants to have a real-time synchronous discussion, Choice: A teacher asks a question and specifies a choice of multiple responses, Database : Enables participants to create, maintain and search a bank of record entries, External tool: Allows participants to interact with LTI compliant learning resources and activities on other web sites, Feedback: For creating and conducting surveys to collect feedback, Forum: Allows participants to have asynchronous discussions, Glossary: Enables participants to create and maintain a list of definitions, like a dictionary, Lesson: For delivering content in flexible ways, Quiz: Allows the teacher to design and set quiz tests, which may be automatically marked and feedback and/or to correct answers shown, SCORM: Enables SCORM packages to be included as course content, Survey: For gathering data from students to help teachers learn about their class and reflect on their own teaching, Wiki: A collection of web pages that anyone can add to or edit and Workshop. After investing shape of data, now these is received data from SQL database converted to excel sheets, the data collected undergoes four pre-processing steps and the data matrix is reduced from 7 columns to main 2 columns shows in table 3.

Table 3. Sample of collected data

Num	Time	IP Add	Stud ID	Name	Action	Statues	Course
1	sun 17-2-2010	127.01.01	209	a	Wiki	Ok	CR217
2	sun 11-3-2010	127.01.02	104	m	forum	Ok	CR217
3	sun 11-3-2010	127.01.03	208	s	Wiki	No	CR217
4	sun 11-3-2010	127.01.04	207	b	messages	Ok	CR217

Second step converts data from textual values to numeric ones in order to deal with identification numbers. Third step converting to ARFF/CSV to us in WEKA shows in next table 4 as sample.

Table 4. Example of data after selecting main segments

Stud. ID	Bus. & Tech. - BIS – English – Courses CR217							
	Actions							
	Wiki	Chat	Message	Quiz	Forum	Feedback	Assignments	Workshop
A209	1	0	0	0	0	1	0	1
M2010	0	0	1	0	1	0	1	0
A210	1	0	1	0	0	0	0	0

Third step apply clustering technique using K-Mean algorithm to support decision for teacher to improve students qualifications. K-mean algorithm cluster students usage to 3 clusters each one of them had some features as follow:

- In fist cluster; group 1 shows no use for assignments, very low for sending, read and replay messages and very few answering quizzes or passed. *So*, cluster these students to non-active student and it had 12% of all data.

- In second cluster; group 2 shows highly quizzes done and passed, highly messages read and replay, low quiz failed and highly total time in assignments. *So,* cluster these students to very active students and it had 42% of all data.

- In third cluster; group 3 shows moderate assignments, reading message and failed in quizzes witch attended. *So,* cluster these students to active students and it had 46% of all data.

Table 6: clustering technique results using K-M

Clustering Technique results using K-M Algorithm		Results
Clusters	**Features**	
G 1 C0=	*No assignments* *Very low messages read and replay* *Very few quiz done and passed*	*Non- Active Students* *12%of over all data*
G 2 C1=	*Highly quiz done and passed*	*Very Active Students*
	Highly messages read and replay *Low quiz failed* *Highly total times in assignments*	*42% of Data*
G 3 C2=	*Moderate assignments* *Moderate reading messages* *Failed at quiz attending*	*Active Students* *46% of Data*

6. CONCLUSIONS

The proposed system HEDMS using higher technologies to solve main problem in education and after investigating the result for proposed system can support teacher with information to make decisions about the students and the Moodle activities of the course in order to improve the students' learning. In this work we have shown a survey concerning the application of data mining in course management systems along with a case study and tutorial about the Moodle system. We have described how different data mining techniques can be used in order to improve the students' learning by supporting decision of teacher.

All these techniques can be applied separately in a same system or together in a hybrid system. Although we have described the most general and well-known data mining techniques, there are as well other specific data mining techniques that are also used in e-learning such as classification, clustering and association rules. This technique is used for data cleansing, spotting emerging trends and recognizing unusually good or bad performers of students. In e-learning, outlier detection can be used for assisting instruction in the detection of learners' irregular learning processes.

7. FUTURE WORK

The future work of this study will be focused on how we could develop the proposed approach and enhancing the approach by using DM classification technique to provide the teacher's best result and support high level of management with a good decision.

8. REFERENCE

[1] Salmon, Gill, 2011. "E-moderating: the key to teaching and learning online", Vol. 45, pp. 49-53.

[2] Karrer, T, 2012, "What is eLearning 2.0: Elearningtech", www.blogspot.com.

[3] Bates, 2011. "Technology, e-Learning and Distance Education London: Routledge".

[4] Mamcenko, J. and Sileikiene, I., Fep. 2006, "Analysis of E-exam Data using Data Mining Techniques", Spain.

[5] Ayesha A., Mustafa T. and Khan M. I., 2010, "Data Mining Model for Higher Education System". European Journal of Scientific Research, Vol. 43, No.1, pp.24-29.

[6] Garcia E., Romero C., Ventura S. and Calders T., 2013, "Drawbacks and solutions of applying association rule mining in learning management systems", Proceedings of the International Workshop on Applying Data Mining in e-Learning, pp. 13-23.

[7] Merceron, A. andYacef, K., 2013, "Interestingness Measures for Association Rules in Educational Data". Proceedings of the first International Conference on Educational Data Mining, Canada, pp. 35-41.

[8] Merceron, A. and Yacef, K., 2009, "Educational Data Mining: a Case Study, School of Information Technologies, University of Sydney", Australia, Vol. 11, No. 5.

[9] Han, J. and Kamber, M., 2001, "Data mining: concepts and techniques", San Francisco, Morgan Kaufman.

[10] Merceron, A. and Yacef, K., 2004. "Mining Student Data Captured from a Web-Based Tutoring Tool: Initial Exploration and Results", Journal of Interactive Learning Research (JILR), Vol.15, No.4, pp. 319-346.

[11] Herin, D., Sala, M. and Pompidor, P., 2002, "Evaluating and Revising Courses from Web Resources Educational", In Int. Conf. on Intelligent Tutoring Systems, Spain, pp. 208–218.

[12] Chen, Han, and Yu. 1996. "Knowledge and Data Engineering Data mining: An overview from a database perspective", IEEE Trans, Vol. 30, No.3, pp. 8:866-883.

[13] Fayyad U., Piatetsky G. and Frawley W. J, 1991. "Press

[14] Gartner, February 2000, "Evolution of data mining", Gartner Group Advanced Technologies and Applications Research Note.

[15] Turban, E., Aronson J., Liang T. and Sharda R, 2007, "Decision Support and Business Intelligence Systems", eighth edition. Prentice Hall.

[16] Hwang, G., "A Knowledge-Based System as an Intelligent Learning Advisor on Computer Networks", 2000, Vol. 30, No.3, pp. 153–158.

[17] Matsui, T., and Okamoto, T., 2003, "Knowledge Discovery from Learning History Data and its Effective Use for Learning Process Assessment under the e-Learning Environment", In: Crawford, Society for Information Technology and Teacher Education International Conference, pp. 3141–3144.

[18] Mostow, J., Beck, J., Cen, H., Cuneo, A., Gouvea, E., and Heiner, C., 2013, "An Educational Data Mining Tool to Browse Tutor-Student Interactions: Time Will tell",in: Proceedings of the Workshop on Educational Data Mining, pp.15–22.

[19] Mohammad H. and Jafar H, September, 2011. "Using Educational Data Mining Methods to Study the Impact of Virtual Classroom in E-Learning", IJDKP.

[20] Gabrijela D. and Dragana P. September, 2011, "The Use of Data Mining Methods For Analyzing And Evaluating Course Quality In The Moodle System", IJDKP, Vol. 47, No.3, pp. 17-27.

[21] Sonali A., Pandey G. and Tiwari M., 2010. "Data Mining In Education: Data Classification and Decision Tree Approach" Vol. 11, No. 5.

[22] Félix C., Alfredo V., Àngela N. and Francisco M., 2010, "Applying Data Mining Techniques to E-Learning Problems" Vol. 2, No. 6.

[23] Mamcenko, J., Sileikiene, I., 2013, "Analysis of E-exam Data using Data Mining Techniques", Vol. 10, No.7.

[24] Khedr, A. Elsayed, "Towards Three Dimensional Analyses for Applying E-Learning Evaluation Model: The Case of E-Learning in Helwan University", International Journal of Computer Science Issues (IJCSI), Vol.9: Issue: 4: No.1, July 2012, Pp.161-166.

[25] Amira M. Idrees, Ayman E. Khedr and Ahmed I. El Seddawy, 2014, "Performance Tuning of K-Mean Clustering Algorithm a Step towards Efficient DSS ", Vol. 2, No. 6.