International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358`

Forecasting Some Aspects of E-Learning Educational System Using Data Mining Techniques

Jyoti Upadhyay[1], Pratima Gautam[2]

[1]Ph.D. Research Scholar, AISECT University Bhopal, (M.P.), India

[2]Dean IT Department, AISECT University Bhopal,(M.P.), India

Abstract: ***Currently in INDIA there is an increasing interest in e-education. Whenever we have a new start that time we have to face lots of problem to manage it, Data mining techniques provides lots of tools for this problem. This work provides the application of data mining to e-learning educational systems, particular web-based courses, well-known learning content management systems, and adaptive and intelligent web-based educational systems. Each of these systems has different tricks,algorithms and objectives for knowledge discovering. After preprocessing the available data in each case, data mining techniques can be applied like statistics and visualization; clustering, classification and outlier detection; association rule mining and pattern mining.***

Keywords: Data Mining, E-learning, EDM, Prediction, Association.

1. Introduction

Nowadays Data mining applications using in e-learning systems is a rapidly growing phenomenon. In data mining and e-learning system predictive, association, classification modeling's are interesting applications on predicting student performance, designing curriculum, and trainer's performance predication are the usefulness of e-learning systems .predict/classify a student's performance is a very important in e-learning environments. A very hopeful arena to manage objective of e-learning is the use of Data Mining (DM) [14]. There are many different objectives in e-learning, such as: grouping student with similar characteristics [16], to group students who are failure-driven and find common misconceptions those students possesses [15], to find drop-out rates [17], etc…. In fact, one of the most useful DM tasks in e-learning is classification. Data mining techniques can potentially help identify the patterns that may be profitable to the e-learning business [9].Figure 1, shows results from the work done in [11] which studies the effect of student waiting times in a tutoring session and the efficiency of the tutoring session. The first part of the graph shows tutoring time is mostly higher than waiting time implying the higher efficiency of the session, the send half of the graph emphasizes that result, which represents the waiting factor in a log scale.

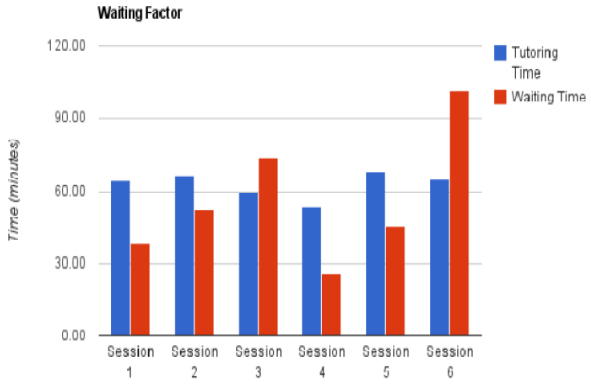

Figure 1. Tutoring time (T_t) and waiting time (T_w) per session

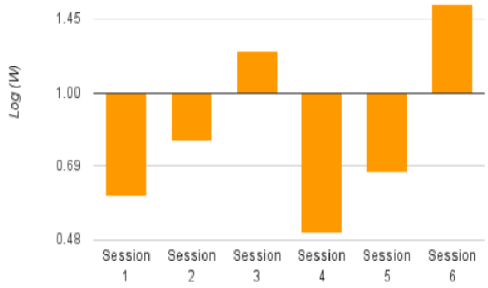

Figure 1: Effective teaching session resulted by data mining (source 11)

1.1Data Mining

Data mining basically type of database analysis that attempts to discover interesting and useful patterns or relationships in huge amount of data. Data mining combines tools from statistics and artificial intelligence like neural networks and machine learning) with database management to analyze large digital collections, which is called data sets/Training set. The analysis uses advanced statistical methods, such as cluster analysis, classification and sometimes employs ARTIFICIAL INTELLIGENCE or NEURAL NETWORK techniques. Han and Kamber (1996) [2] describes data mining software that allow the users to analyze data from different dimensions, categorize it and summarize the relationships which are identified during the mining process.

The steps involve in extracting knowledge from data are: shown in Figure 1.A major goal of data mining is to discover hide and previously unknown relationships from the data set, especially when the data comes from different databases. Data mining useful in many fields, including retail sales, bioinformatics, higher education, counter-terrorism etc. In recent years, there has been increasing interest in the use of data mining to investigate scientific questions within educational research, an area of inquiry termed educational data mining.

1.2 Educational Data Mining

In recent years, there has been increasing interest in the use of data mining to investigate knowledge within educational research, this inquiry area known as educational data mining. **Educational Data Mining** (EDM) is concerned with develop some methods which can be helpful in exploring the unique types of data that come from educational atmosphere, and using those methods to better understand students, and the settings which they learn in. from statistics, machine learning, and data mining, EDM develops methods and applies techniques to analyze data which are collected during teaching and learning. Educational data mining is a suite of computational and psychological methods for understanding how students learning behavior.EDM is concerned with developing methods to explore the unique types of data in educational settings and, using these methods, to better understand students and the settings in which they learn [5].

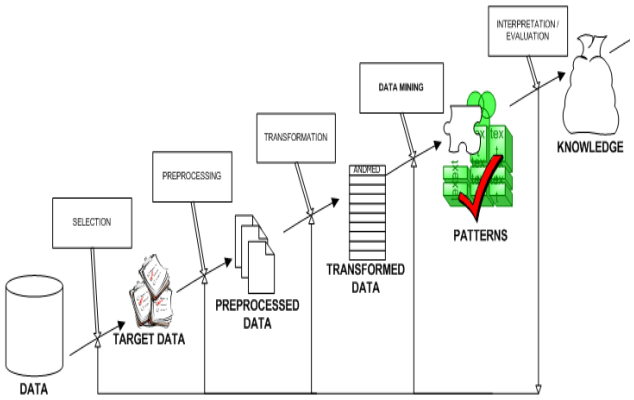

Figure 2: steps for extracting knowledge

1.3 E-learning Educational System

E-learning referred to as web based education. Many organizations worldwide are now delivering not only education over the Internet but training also. Some college and university offer courses and degree program through e-learning. Some e-learning companies or institutions offer online tutoring. They offer courses at specific grade levels. Increasingly, training and support for teachers is occurring online, and a number of institutions now offer either partial or complete secondary diplomas through e-learning. According to **Megha Banduni** reports at express computer "e-learning permits the delivery of knowledge and information to learners at an accelerated pace, opening up new vistas of knowledge transfer. Early adopters are companies that have tried to supplement face-to-face meetings, demonstrations, training classes and lectures with this technology. "The adoption of e-learning in all spheres—corporate, schools, universities, etc—is low at present. According to **Shimaa Abd in E-learning using data mining define E-learning** as a new context for education where large amounts of information describing the continuum of the teaching-learning interactions are endlessly generated and ubiquitously available".

2. Aspects of e-Learning

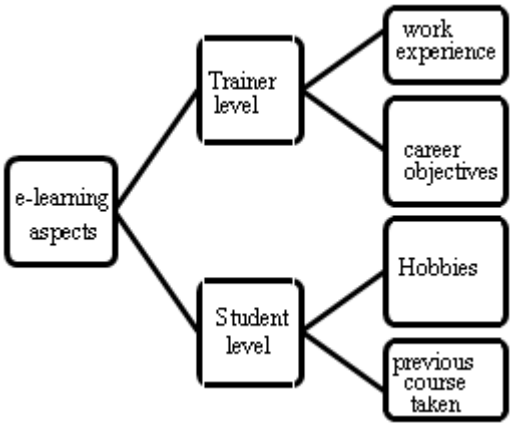

Figure 3: Aspects of E-Learning

We can categorize the e-learning field mainly in two groups of users: one is the learners and the learning providers. Figure: 2 shows this categorization .Private training companies, governmental organizations and local authorities play role of learning provider. They are providing training to make their courses access available on internet. Learning provider /Companies maintain database for learners to store all personal details like name, age, gender, address, postcode, and educational-relevant details such as qualifications. Moreover having information like work experience, career objectives, income range, previous courses taken and courses of interest would be of great value to be able to predict future behavior of different classes of employed professional people. Also other information such as personal interests and hobbies would be very valuable for data mining tool in order to discover hidden patterns by building intelligent models based on the huge amount of data [4].The main components of a learning atmosphere may be the teacher, the learner, the content, the assessment mechanism, communication and collaboration mechanisms, and the administrative aspects. In e-learning some of them changes in style according to new medium. In e-learning environment, the students have much richer sources of information than the traditional instructor-the internet resources. These data can be analyzed from various levels and perspectives and analysis can shows different aspects. In some level like individual course we can consider participation in lectures, assignments, enrolling in midterm and final exams. As traditional tasks like overall student performance or drop out prediction, it becomes possible to track with the help of data mining. How different learning resources like video /audio lectures, handouts, wikis, hypermedia, quizzes are useful. Students do progress or not with assignments self-assessment test and questionnaires, etc can also examine.

3. Data mining Methodology used in e-learning

Many authors establish the research opportunities in e-learning and data mining on the basis of some technique of data mining. In [6], studies on how Data mining techniques can improve learning capability in e-learning environments successfully. In [3], data clustering technique was suggested for student diagnosis and promote group-based collaborative learning. in [1] authors presented possibilities of Web Mining techniques to face some challenges in distance education. When we talk about place of data mining in e-learning we found that many areas of DM helps to improve e-learning as a quality education technique. Using Classification Models of DM we can forecast student growth, student behavior, trainers' performance etc. Through Clustering We can build a model to improve learning process. We are trying to explain these techniques in this paper as we mention in Introduction section also.

3.1 Clustering Method

Clustering is a group of clusters. Clustering is a division of data into groups of similar objects. Each group, called cluster, consists of objects that are similar between themselves and dissimilar to objects of other groups. The goal of clustering is descriptive where as classification is predictive (Veyssieres and plant, 1998). So through clustering we can find new description of a data set from existing one. A research on how Data Mining techniques could be successfully improves the learning processes in e-learning environment was presented in [6]. Here, data clustering is suggested as a means to promote group based collaborative learning and to provide incremental student diagnosis. Some researchers presented the use of clustering techniques to group similar course materials which may help e-learning users to find and organize distributed courseware resources. To group similar learning materials, an element of clustering tool used in the implementation of the Bisection K-Means algorithm. In [8] Kohonen's well-known SOM algorithm was used to devise an intelligent searching tool to cluster similar learning material into classes. In [7] Clustering tools was proposed to group similar learning credentials and it should be based on their domain.

3.2 Classification and/ or prediction Method

There are two forms of data analysis that can be used for predict future data trends from a training set. These are Classification and Prediction. Classification predicts categorical and prediction models predict continuous valued functions. Figure 4 shows classification process which source is zeepidea.com .There are many method used in data mining for classification like fuzzy logic, Artificial Neural Networking ,Decision tree etc.

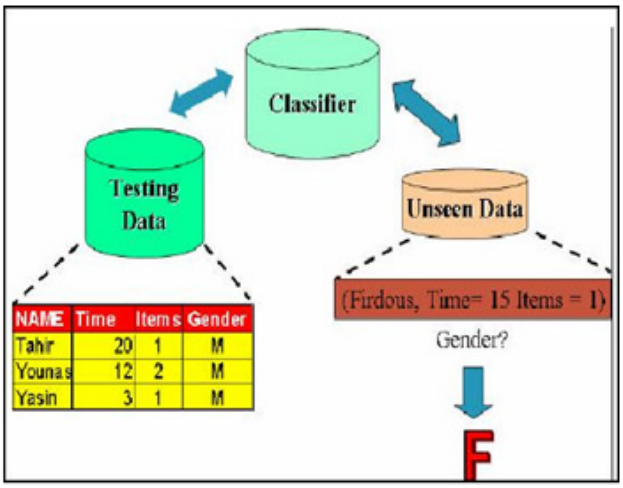

Figure 4: Classification Process- Model Construction

By Mizue, K in [10] put a navigation support system which is based on an Artificial Neural Network to take decision on the appropriate navigation strategies. The Neural Network was used as a navigation strategy decision module in the system. a neural network is a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes. According to the DARPA Neural Network Study (1988, AFCEA International Press, p. 60): In fuzzy method a t mining A Two-Phase Fuzzy Mining and e-learning algorithm was described in [12]. Tsai, C.J., Tseng describe an association rule mining algorithm, called *Apriori*, with fuzzy set theory to find embedded information. This information can be used in feedback process of trainer/faculties to refining performance. In a second phase, it uses an inductive e-learning algorithm to find the concept descriptions which indicates the missing concepts during students' learning. Behrouz Minaei-Bidgoli and others in[13] presents an approach to classifying students in order to predict their final grade based on features extracted from logged data in an education web-based system. We design, implement, and evaluate a series of pattern classifiers and compare their performance on an online course dataset. A combination of multiple classifiers leads to a significant improvement in classification performance. Furthermore, by learning an appropriate weighting of the features used via a genetic algorithm (GA).

3.3 Association rule Method

User can withdrawal interesting patterns automatically through Data mining) from large data collections [18]. Association rules mining discovers relationships among attributes in databases, producing if-then statements concerning attribute-values [19]. Enrique García and others in "Drawbacks and solutions of applying association rule Mining in learning management systems" explain that an association rule X ->Y expresses that in those transactions in the database where X occurs; there is a high probability of having Y as well. X and Y are called respectively the antecedent and consequent of the rule. The strength of such a rule is measured by its support and confidence. The confidence of the rule is the percentage of transactions with X in the database that contain the consequent Y also.

Association rule mining has been applied to e-learning systems for finding correlations between items in a dataset, including the following tasks: building recommender agents for on-line learning activities [20], automatically generate and recommend learning materials [21], finding students' mistakes that are usually occurring [25], discovering interesting relationships from student's usage information to provide feedback to instructor [23],identifying attributes patterns of performance difference between various groups of students [22], finding out the pattern of learner's behavior [24], guiding the search for best fitting transfer model of student learning [26] etc.

Apart from above mention method in e-learning we can use Statistical analysis for statistical data, Bayesian Network for prediction, Social Network Analysis for Visualization and clustering. Other data-mining techniques that can be used in e-learning are prediction techniques, visualization techniques, and case-based reasoning.

4. Conclusion

In this work we have shown how useful the application of data mining techniques in e-learning educational system to forecast course management systems, student behavior, learner performance etc.. Although we have shown these techniques separately, they can also be applied together in order to obtain interesting information in a more efficient and faster way. The application of data mining in e-learning educational systems has specific requirements which are not present in other domains; requirements need to take into account of educational aspects. In this paper, we are trying some detail of current state of the research in Data Mining applied to e-learning, highlighting its future perspectives and opportunities, as well as its limitations. If India adopt e-learning education system in full flash manner then we are sure through data mining technique we mange it in a better way.

References

[1] Ha, S.H., Bae, S.M., Park, S.C.: Web Mining for Distance Education. In: IEEE International Conference on Management of Innovation and Technology, ICMIT'. (2000).
[2] Jiawei Han, Micheline Kamber, Data Mining: Concepts and Techniques, 2nd edition, (2006).
[3] Tang, T.Y., McCalla,Smart Recommendation for an Evolving e-Learning System:Architecture and Experiment. International Journal on e-Learning (2005)
[4] Margo Hanna, "Data mining in the e-learning domain", Campus-Wide Information Systems, (2004)
[5] Baker, R, Data Mining for Education. To appear in McGaw, B., Peterson, P., Baker, E. (Eds.) International Encyclopedia of education (3rd edition). Oxford, UK: Elsevier, (2010).
[6] Margo, H.: Data Mining in the e-Learning Domain. Computers & Education (2004)
[7] Hammouda, K., Kamel, M.: Data Mining in e-Learning. In: Pierre, S. (ed.): e-Learning Networked Environments and Architectures: A Knowledge Processing Perspective. Springer-Verlag, Berlin Heidelberg New York (2005).
[8] Drigas, A., Vrettaros, J.: An Intelligent Tool for Building e-Learning Contend-Material Using Natural Language in Digital Libraries. WSEAS Transactions on Information Science and Applications (2004)
[9] Hanna, M.. Data mining in the e-learning domain. *In Campus-Wide Information Systems,* Volume21, Number 1,(2014)
[10] Mizue, K., Toshio, O.: N3: Neural Network Navigation Support-Knowledge-Navigation in Hyperspace: The Sub-symbolic Approach. Journal of Educational Multimedia and Hypermedia 10(1) (2001)
[11] Rojas, I.G.; Garcia, R.M.C, "Towards Efficient Provision of Feedback Supported by Learning Analytics,"*Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on*.(2012)
[12] Tsai, C.J., Tseng, S.S., Lin, C.Y.: A Two-Phase Fuzzy Mining and Learning Algorithm for Adaptive Learning Environment. In: Alexandrov, V.N., et al. (eds.): International Conference on Computational Science, ICCS 2001. LNCS Vol. 2074. Springer-Verlag, Berlin Heidelberg New York (2001)
[13] Behrouz Minaei-Bidgoli , Deborah A. Kashy , Gerd Kortemeyer , William F. Punch: Predicting student performance: an application of data Mining methods with the educational web-based system LON-CAPA, 33rd ASEE/IEEE Frontiers in education Conference November 5-8, (2003)
[14] Romero, C., Ventura, S. Educational Data Mining: a Survey from 1995 to 2005.Expert Systems with Applications, (2007)
[15] Yudelson, M.V., Medvedeva, O., Legowski, E., Castine, M., Jukic, D., Rebecca, C. Mining Student Learning Data to Develop High Level Pedagogic Strategy in a MedicalITS. AAAI Workshop on Educational Data Mining, (2006)
[16] Chen, G., Liu, C., Ou, K., Liu, B. Discovering Decision Knowledge from Web Log Portfolio for Managing Classroom Processes by Applying Decision Tree and Data CubeTechnology. Journal of Educational Computing Research (2000),
[17] Cocea, M., Weibelzahl, S. Can Log Files Analysis Estimate Learners' Level of Motivation? Workshop on Adaptivity and User Modeling in Interactive Systems, Hildesheim, (2006)
[18] Klosgen, W., & Zytkow, J.: Handbook of data mining and knowledge discovery. Oxford University Press, New York (2002).
[19] Agrawal R., Imielinski, T., Swami, A.N.: Mining Association Rules between Sets of Items in Large Databases. In: Proc. of SIGMOD (1993) .
[20] Zaïane, O.: Building a Recommender Agent for e-Learning Systems. In: Proc. of the Int. Conf. in Education (2002).
[21] Lu, J.: Personalized e-learning material recommender system. In: Proc. of the Int. Conf. on Information Technology for Application (2004)
[22] Minaei-Bidgoli, B., Tan, P., Punch, W.: Mining interesting contrast rules for a web-based educational system. In: Proc. of the Int. Conf. on Machine Learning Applications (2004)
[23] Romero, C., Ventura, S., Bra, P. D.: Knowledge discovery with genetic programming for providing feedback to courseware author. User Modeling and User-

Adapted Interaction: The Journal of Personalization Research, (2004).
[24] Yu, P., Own, C., Lin, L.: On learning behavior analysis of web based interactive environment. In: Proc. of the Int. Conf. on Implementing Curricular Change in Engineering Education (2001) .
[25] Merceron, A., & Yacef, K.: Mining student data captured from a web-based tutoring tool.Journal of Interactive Learning Research, (2004) .
[26] Freyberger, J., Heffernan, N., Ruiz, C.: Using association rules to guide a search for best fitting transfer models of student learning. In: Workshop on Analyzing Student-Tutor Interactions Logs to Improve Educational Outcomes at ITS Conference (2004)

Author Profile

Mrs Jyoti Upadhyay is Assistant Professor in the Department of Computer Science, SSSSMV, Hudco, Bhilai, Chhattisgarh, India. Obtained her M.C.A degree from Govt Engineering college Raipur, Chhattisgarh, India (NIT Raipur)(2003). She is currently doing research in Data Mining and Knowledge Discovery. She has published one international paper.

Dr. Mrs. Pratima Gautam received Ph.D. degree from MNIT Bhopal, M.P.(2012). At present, she is working as Dean of Department of IT in AISECT University Bhopal, M.P. Chhattisgarh, India. She has authored a commendable number of research papers in international/national Conference/journals. His research interests include E-learning, Data Mining, and soft computing.