



Predict Student improvement by using Data Mining Techniques and Component

M. N. Quadri

Assistant Professor, Department of Computer Science

Nilkanthrao Shinde Science & Arts College

Bhadrawati, Dist. Chandrapur

quadrimn@gmail.com

Abstract

Data mining in Education is one of the interdisciplinary research areas which are used to identify the methods and explore data in education field. Educational Data Mining (EDM) is to study related questions of education. It gives the planning for enhancement of teaching and learning process. In this paper I have study and focuses on methods which are used for educational data to built models for improving performance in academic by the learner. The study shows the prediction of student for improvement in studies using data mining algorithms.

Keywords- *Educational data mining (EDM), EDM Tools, Data Mining Techniques, Student Performance, Prediction.*

I. INTRODUCTION

Educational Data mining (EDM) is a technique which is applies for educational data. It is gives statistical, machine learning and data mining (DM). It is used to analyze data in order to resolve research issues [1]. EDM is used for better understanding of student's data [2]. EDM is one of the emerged research area for researchers in the world which include offline education, E-learning and Intelligent tutoring system [3]. EDM used for some quality of edification and efficacious cognition process [4]. The expected growth difficulties arise in application of EDM such as different types of data, varied objectives and techniques. The different researchers having view on EDM objectives such as student modeling, learning system, domain modeling, building the computational models and study the effects of resources [5][6]. In this paper I have try to evaluating faculty performance to be used with different classification algorithms that predicts faculty performance. The outcome show that can predict the result of the faculty and then it becomes possible for taking required action [7]. The data mining techniques can be used to avail teachers to manage their classes, understand their students; learning processes as well as the feedback to learners [8][5].

The objective of this survey paper is to review, different Data Mining Methods applied to EDM related. This survey paper presents a review of most of the types of Data Mining methodologies applied to EDM till date. In this survey the most recent works in education that has been resolved by utilizing DM techniques and identifies and suggests few research opportunities and future scopes in EDM.

II. LITERATURE REVIEW

EDM is used for analyze educational data with an objective to study edifying questions. The most germane studies introduces EDM and describes the different groups of utilize, types of scholastic environments, and the data they provide [1]. The applications [9] of data mining in educational institution to extract useful information from the large data sets and providing analytical tool to view and use this information for decision making processes by taking real life examples. The Data Mining algorithms [5] can pick pedagogically important mines contained in the data stores obtained from the educational system. These knowledge help to manage classes understand students and practice it on their teaching and to support learner reflection and provide proactive feedback to learners. EDM [10] helps to learn and develop models for the growth of education environment. It provides decision makers a better understanding of student learning and the environment setting in as of EDM. Difficulties in learning programming order to initiate development of learning materials for elementary programming courses [11]. The survey provides information of the hurdles felt, experienced and perceived when learning and teaching programming.

The different parameters [7] utilized in evaluating faculty performance to be utilized with different relegation algorithms that prognosticates faculty performance utilizing the advantages of Distributed Data Mining. It is shown that [12] proposed an educational data mining model for predicting student performance in programming courses. The proposed model includes three phases; data pre-processing, attribute selection and rule extraction algorithm. A students' performance prediction system using Multi Agent Data Mining is proposed by [13], to predict the performance of the students based on their data with high precision and provide the weaker students by optimization rules.

Adopted a data mining approach [14] applied to discover students' performance models in supervised and unsupervised assessment tools of a course in an engineering degree program.

III DATA MINING PROCESS

The data mining is a process in which user have to identifying valid, novel, unknown, potentially utilizable information and understandable patterns from data in database [15][16]. Data mining [13] has different steps as shown in following figure 1.

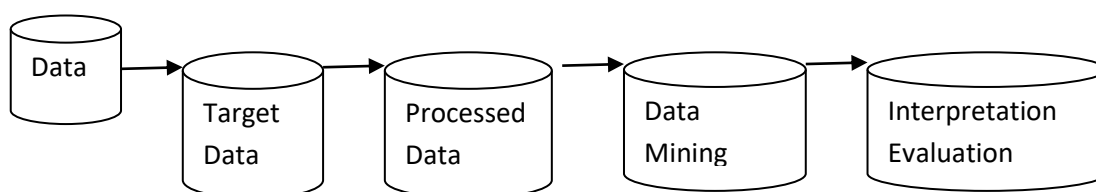


Figure 1. Data Mining Process



- a) Data Amassing and Processing: in this step data collection is executed. Predicated on our goals, data can be amassed from different environments or offices, such as banks, schools, markets, educational environments and etc. After accumulating data, a sole data warehouse may be providing for keeping this data for subsequent processing steps. The amassed data further will be processed and all faults, redundant data and etc., will be abstracted.
- b) Data Transformation: in this step predicated on utilization or implements which we will utilize for analyzing data, we require to transform data. For example for utilizing with Weka we require to make .csv files and etc.
- c) Pattern Revelation: in this step with applying data mining techniques such as clustering, classification, and etc., we will be endeavor to discover pattern from that data.
- d) Knowledge revelation: in this step we will be endeavour to utilize the extracted patterns for more examination or extracting association rules or further analyses.
- e) Evaluation: in this step, with testing our extracted cognizance, the percentage of efficiency of that knowledge will be declared.
- f) Action: Determinately with discovering all impuissance or efficiency of this extracted cognizance we can utilize this knowledge for sundry usages or applications.

IV USERS, TOOLS AND TECHNIQUES OF DATA MINING

Different people are involved with educational data mining of which there are four main users such as

- i. Learners - Learners are interested in understanding student needs and methods to improve the learner's experience and performance.
- ii. Educators - Educators attempt to perceive the learning process and the methods they can use to improve their teaching methods. Educators can utilize extensibility of EDM to determine how to organize and structure the curriculum, the best methods to present course information and the tools to use to engage their learners for optimal learning outcomes.
- iii. Researchers - Researchers fixate on the development and the evaluation of data mining techniques for efficacy.
- iv. Administrators - Administrators are responsible for allocating the resources for implementation in institutions.

DM tools are normally designed more for power and flexibility than for simplicity [1]. Due to the rapid magnification of educational data, there is a desideratum to summarize the tools according to their function/features, integrated techniques and working platforms. MINEL, LOCO, EPRules, GISMO, O3R, Synergo/ColAT, LISTEN Mining tool, CIECoF, PDinamet, Meerkat, MMT tool, TADAED are examples of EDM tools

There are several applications or tasks in teaching environments that are resolved through DM. [17][18] suggests four key areas of application for EDM: ameliorating student models, amending domain models, finding out the education support provided by learning software system, and research project into learning and learners.



There are many techniques in EDM to meet and achieve its objectives. Very few of them are categorized into three and are follows:

1. Prediction: This technique is utilized to derive predicated variable (single variable) from predictor variables It has three types.

A. Classification: used to presage class label from (discrete or perpetuate). Some popular classification methods possess logistic regression, SVM and decision trees.

B. Regression: used to presage from continuous variable. Some of the famous regression methods of educational data mining include linear regression, neural networks.

C. Density Estimation: probability density function is utilized to predicted variable. Density estimator can be predicated on variety of kernel functions, considering Gaussian function also.

2. Clustering: Clustering is an unsupervised relegation process. It is utilized for grouping objects into classes of kindred objects. Data items are partitioned into groups or subsets (clusters) predicated on their neighborhood and connectivity within N-dimensional space. In educational data mining, it uses clustering to group students according to their cognition.

3. Relationship mining: Relationship mining is utilized to determine relationship between variables in a data set and form rules for categorical purport. Relationship mining is relegated into four types:

A. Association rule mining: This method is utilized to identify relationship between attributes in data set, extracting intriguing correlations, frequent patterns among data items for finding students' mistakes most often recrudescence together while solving exercises

B. Correlation mining: This method is utilized to find Linear correlations between variables (positive or Negative).

C. Correlation analysis: It is utilized to find the most vigorously correlation attributes.

D. Sequential pattern mining: This method is specifically utilized to find inter-session patterns such as the presence of a set of items followed by another items in a time-coherent set of sessions or episodes predicated on temporal relationship between variables to soothsay which group a learner belongs to.

E. Causal data mining: This method is utilized to find causal relationship between variables by analyzing the covariance of two events or by utilizing information about how one of the events was trigger.

V. CONCLUSION

The aim of this paper gives a complete survey towards the research papers which would have discussed different Data Mining Methods and algorithms applied to EDM context. These surveys are very helpful for educational data mining methods and tools which is used to improvements in teaching and predicting the performance of Students so that to early prediction of the learner.

REFERENCES

- [1] Cristobel Romero, Sebastian Ventura "educational data mining: a review of the state of the art" iee transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 40, no. 6, november 2010



- [2] Mrinal Pandey, Vivek Kumar Sharma, PhD., "A Decision Tree Algorithm Pertaining to the Student Performance Analysis and Prediction", International Journal of Computer Applications (0975 – 8887) Volume 61– No.13, January 2013
- [3] F. A. Saeed, "Comparing and Evaluating Open Source E-learning Platforms", International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307, Volume-3, Issue-3, 224-249, July 2013.
- [4] C. Romero, S. Ventura, and P. De Bra, "Knowledge discovery with genetic programming for providing feedback to courseware author," UserModel. User-Adapted Interaction: J. Personalization Res., vol. 14, no.5, p. 425-464, 2004.
- [5] Agathe MERCERON, and Kalina YACEF, "Educational Data Mining: a Case Study ", Proceedings of the 12th international Conference on Artificial Intelligence in Education AIED 2005.
- [6] O.S. Akinola, B.O. Akinkunmi, T.S. Alo , " A Data Mining Model for Predicting Computer Programming Proficiency of Computer Science Undergraduate Students", African Journal of Computing & ICT January, 2012, Vol 5. No. 1 - ISSN 2006-1781
- [7] Priyanka r shah, prof. Dinesh b vaghela , Dr. Priyanka sharma , "predicting and analysing faculty performance using distributed data mining", international journal of emerging technologies and applications in engineering, technology and sciences (ij-eta-ets) issn: 0974-3588, December 2014
- [8] <http://computation.llnl.gov/casc/sapphire/overview/overview.html>
- [9] V. Sarala, Dr. v.v. Jaya Rama krishnaiah, "empirical study of data mining techniques in education system", international journal of advances in computer science and technology (ijacst), vol. 4 no.1, pages: 15 –21 2015
- [10] Jasvinder Kumar, A Comprehensive Study of Educational Data Mining, International Journal of Electrical Electronics & Computer Science Engineering Special Issue - TeLMISR 2015, ISSN: 2348-2273
- [11] Essi Lahtinen, Kirsti Ala-Mutka, Hannu-Matti Järvinen, "A Study of the Difficulties of Novice Programmers", ACM ITiCSE, June 27–29, 2005
- [12] A.F.ElGamal, "An Educational Data Mining Model for Predicting Student Performance in Programming Course", International Journal of Computer Applications (0975 – 8887) Volume 70– No.17, May 2013
- [13] Dr. Abdullah AL-Malaise, Dr. Areej Malibari, and Mona Alkhozae, "Students' Performance prediction System Using multiagent Data mining Technique ", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.4, No.5, September 2014
- [14] Anwar M. A., Naseer Ahmed, "Information Mining in Assessment Data of Students' Performance", ISSN: 2319-5967 ISO 9001:2008 Certified International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 1, Issue 2, November 2012
- [15] Sweta Rai, Ajit Kumar Jain, "Students' Dropout Risk Assessment in Undergraduate Courses of ICT at Residential University – A Case Study", International Journal of Computer Applications (0975 – 8887) Volume 84 – No 14, December 2013
- [16] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, R "The KDD Process for Extracting Useful Knowledge from Volumes of Data," Communications of the ACM, (39:11), pp.27-34, (1996).



- [17] R. Baker and K. Yacef, "The state of educational data mining in 2009: A review and future visions," J. Educ. Data Mining, vol. 1, no. 1, pp. 3–17, 2009.
- [18] R. Baker, "Data mining for education," in International Encyclopedia of Education, B. McGaw, P. Peterson, and E. Baker, Eds., 3rd ed. Oxford, U.K.: Elsevier, 2010.