



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirce.com

Vol. 6, Issue 6, June 2018

Prominent Tasks in Educational Data Mining

Kapila Kundu

Research Scholar, Department of Computer Science and Engineering, GJUS&T University, Hisar, India

ABSTRACT: Educational Data Mining (EDM) is a prominent area that includes various methods, tasks and models which are used to explore the data extracted from educational systems. It have various technique used on different type of data. These techniques are used to assess the behavior, performance of the student and also provide required feedback to the student and the teacher. This paper reviews the useful studies carried out in educational data mining task. This paper has been categorized on the basis of EDM tasks by different researchers. The main focus of this review has been on the availability of data, experimental work, conclusion, and future directions. The conclusion is provided in the last section.

KEYWORDS: Data mining, Educational data mining, Prediction, Classification, regression.

I. INTRODUCTION

By Educational data mining website, Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in. Educational data mining is a process which converts educational data into useful information that can greatly influence research in the field of education. EDM has emerged as an important research field in the last few years. EDM is concerned with developing methods, tasks, models and algorithms for exploring information mined from educational systems [3].

EDM include steps as the basic data mining process like pre-processing, data mining and post-processing, and also applies typical data mining techniques such as classification, clustering, association rule mining, sequential mining and text mining. It also includes other approaches such as regression, correlation, visualization, etc., which are not frequently used in data mining [3]. Educational Data Mining researchers study a variety of areas, including individual learning from educational software, computer supported collaborative learning, computer-adaptive testing, and the factors that are associated with student failure or non-retention in courses. Other key areas of application have been in the improvement of student models, discovering or improving models of a domains' knowledge structure, studying pedagogical support, and looking for empirical evidence to refine and extend educational theories [4, 13].

The publications in Educational Data Mining have grown exponentially in recent years. The first workshop on EDM makes its appearance in 2005. Then EDM emerges in first International Conference on Educational Data Mining in 2008 organized by International Working Group. The first peer-reviewed journal JEDM (Journal of Educational Data Mining), appears in 2009 and two remarkable books on EDM edited by Romero & Ventura entitled: "*Data Mining in E-learning*" [5] and "*The Handbook of Educational Data Mining*" [6] co-edited with Baker, Romero, Ventura, and Pechenizkiy got published in 2006 and 2010 respectively. A great variety of methods have been presented by the researchers in the field of educational data mining through the last several years. Some methods are similar to basic data mining methods, which are used in other domains also. On other hand, some methods are unique to educational data mining. With the reference of literature, methods can be divided into four major classes, which are frequently used in EDM. These are as follows: 1) prediction model; 2) structure discovery; 3) relationship mining; and 4) discovery with models [9]. In this review, the focus is on predictive modeling [13].



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirce.com

Vol. 6, Issue 6, June 2018

II. EDM TASKS

There are numerous applications or tasks that have been employed in educational environment. Many researchers describe variety of functionalities related to student, teacher and organization. These tasks are broadly classified by the gems of educational data mining.

II.1. TASKS BY ALEJANDRO

Alejandro (2014) suggests six types of functionalities for EDM: student modeling, student behavior modeling, student performance modeling, assessment, student support and feedback, and curriculum, domain knowledge, sequencing, and teacher support [9].

- Student modeling

Student modeling is used to classify the learner according to his/her characteristics such as emotion, cognition, domain knowledge, learning strategies, achievements, features, learning preferences, skill, evaluation and affects. The objective of student modeling is to adapt the learner and its environment to fulfill specific learning requirements of individuals [9].

- Predicting student behavior

The main aim of predicting student behavior is to mine particular pattern behavior in order to provide appropriate help to the user. There are many behavior traits such as: gambling, guessing, sleeping, inquiring, requesting help, willingness to collaborate, time series of access and response [9]. The another objective of predicting student behavior is to detect those students who have some type of problem or unusual behavior such as: erroneous actions, low motivation, playing games, misuse, cheating, dropping out, and academic failure [3].

- Detecting student performance

One of the favorite targets of Educational data mining (EDM) approach is to represent and anticipate performance through student modeling. Many indicators help to detect the students' performance. These are as follows: efficiency, evaluation, achievement, competence, resource consuming, elapsed time, correctness, deficiencies, etc [9].

- Assessment

Main focus of assessment is to evaluate the learners' domain knowledge, acquisition, skill development, and achieved outcomes, as well as reflection, inquiring, sentiments. It is required to differentiate student proficiency at the finer grained level through static and dynamic testing as well as online and offline estimation [3,9].

- Student support and feedback

During the interaction between learner and system, the student support given by the computerized educational system is relevant to enhance the learners' performance and achievements, or to correct their misconceptions, bugs, and faults. Pre-emptive and post-failure are two type of penalties. Most of the educational systems should offer functional systems should offer functionalities to track students' feedback with the purpose to express: suggestion, complaints, requests, and evaluations [9].

- Teacher support, domain knowledge and curriculum

Curriculum is an important issue in academics and teachers to develop before delivering instruction to their pupils. A lot of time and efforts are used in authoring, seeking, adapting, and sequencing content resources to employ the curriculum. Academics are involved in the customization of curriculum and teaching practices with the aim at facilitating learners the acquisition of domain knowledge. Domain knowledge is collection of cognitive models of the knowledge components and skills, which are used to train the pupil. Sequencing scheme are used to deliver the contents and curriculum. Teachers support represents the services devoted to facilitate the ordinary work performed by academics, such as: monitoring students, content searching, collaboration, and teachers modeling [9].

II.2. TASKS BY BAKER

Baker (2012) advocates four key areas of EDM tasks: improving student models, improving domain models, studying the pedagogical support provided by learning software, scientific research into learning and learner; and five approaches: prediction, clustering, relationship mining, distillation of data for human judgment and discovery with model [4].

- Improving Student Model



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirce.com

Vol. 6, Issue 6, June 2018

Student models represent information about a student's characteristics or state, such as the student's current knowledge, motivation, meta-cognition, and attitudes. Modeling student individual differences in these areas enables software to respond to those individual differences, significantly improving student learning [10].

- Improving Domain model

It includes the domain of knowledge structure. Domain of knowledge structure means the relationship between the domain attribute. Many researcher developed algorithms that can transform structure directly from the data.

- Pedagogical Support

This area has been studying pedagogical support given by online software, collaborative learning behavior and other domain and also discovering which types of pedagogical support are most effective, either overall or for different groups of students or in different situations [4].

II.3. TASKS BY ROMERO

Romero (2010) enlists advice many more application area for EDM: Analysis and visualization of data, providing feedback for supporting instructors, recommendations for students, predicting student performance, student modeling, detecting undesirable student behaviors, grouping students, social network analysis, developing concept maps, constructing courseware and planning and scheduling[11].

- Analysis and visualization of data

The objective of the analysis and visualization of data is to examine useful information and support decision making. In the educational environment, it can help educators and course administrators to analyze the students' course activities and usage information to get a general view of a student's learning. Statistics and visualization information are the two main techniques that have been most widely used for this task [6].

- providing feedback for supporting instructors

The main objective is to provide feedback to support course instructor in decision making. It also help to improve students' learning, organize instructional resources more efficiently and enable them to take appropriate remedial action. They also stated that this task is different from analysis and visualisation task. Moreover, providing feedback divulges completely new, hidden, and interesting information found in data [6].

- recommendations for students

According to Romero objective of this task is to be able to make recommendations directly to the students with respect to their personalized activities, links to visits, the next task or problem to be done, etc., and also to be able to adapt learning contents, interfaces, and sequences to each particular student. Several DM techniques have been used for this task, but the most common are association-rule mining, clustering, and sequential pattern mining [6].

- predicting student performance

In this task they stated that objective of prediction is to estimate the unknown value of a variable that describes the student. In education, the values normally predicted are performance, knowledge, score, or mark. This value can be numerical/continuous value (regression task) or categorical/discrete value (classification task) [6].

- student modeling

Student modeling include characteristics of student, their skills and declarative knowledge. It encompasses various technique used to detect the flaws in the student learning.

- detecting undesirable student behaviors

The objective of detecting undesirable student behaviour is to detect those students who have some type of problem or unusual behaviour such as: erroneous actions, low motivation, playing games, misuse, cheating, dropping out, academic failure, etc. Several DM techniques mainly, classification, and clustering have been used to reveal these types of students in order to provide them with appropriate help in plenty of time [6].

- Grouping students

The objective is to create groups of students according to their customized features, personal characteristics, etc. Then, the clusters/groups of students obtained can be used by the instructor/developer to build a personalized learning system, to promote effective group learning, to provide adaptive contents, etc. The DM techniques used in this task are classification and clustering [6].

- social network analysis

Social networks analysis (SNA), or structural analysis, aims at studying relationships between individuals, instead of individual attributes or properties. A social network is considered to be a group of people, an organization or social



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirce.com

Vol. 6, Issue 6, June 2018

individuals who are connected by social relationships like friendship, cooperative relations, or informative exchange [6].

- developing concept maps constructing

The objective of constructing concept maps is to help instructors/ educators in the automatic process of developing/ constructing concept maps. A concept map is a conceptual graph that shows relationships between concepts and expresses the hierarchal structure of knowledge [12]. Some DM techniques mainly association rules mining and text mining have been used to construct concept maps [6]

- constructing courseware

The objective of constructing courseware is to help instructors and developers to carry out the construction/development process of courseware and learning contents automatically. On the other hand, it also tries to promote the reuse/exchange of existing learning resources among different users and systems [6].

- planning and scheduling

The objective of planning and scheduling is to enhance the traditional educational process by planning future courses, helping with student course scheduling, planning resource allocation, helping in the admission and counselling processes, developing curriculum, etc. Different DM techniques have been used for this task like association rules mining [6].

III. CONCLUSION

Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in. Educational data mining is a process which converts educational data into useful information that can greatly influence research in the field of education. It also includes e-learning, adaptive hypermedia, intelligent tutoring systems, web mining, etc. The main focus of literature review is on the tasks which are broadly classified by the researcher in the area of educational data mining. Most commonly used tasks are predicting student performance, detecting student behavior and feedback and support. The area related to student characteristics like emotion, cognition, etc., have been described under student modeling. Student performance task involves methods to improve the correctness, efficiency, and achievement of a learner. The most common EDM approaches are oriented to student modeling and assessment tasks.

REFERENCES

- [1]. J. Han, M. Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, 2006.
- [2]. C. Romero, & S. Ventura, "Educational data mining: a survey from 1995 to 2005", *Expert Systems with Applications*, vol. 33, issue 1, pp. 135–146, 2007.
- [3]. C. Romero, & S. Ventura, "Educational data mining: a review of the state of the art", *IEEE Transactions on systems, man, and cybernetics, part C: applications and reviews*, vol. 40, issue 6, pp. 601–618, 2010.
- [4]. R. S. J. D. Baker, & K. Yacef, "The state of educational data mining in 2009: A review and future vision", *Journal of Educational Data Mining*, vol. 1, issue 1, pp. 1–15, 2009.
- [5]. C. Romero, & S. Ventura, "Data mining in e-learning", book series 4, WIT Press, 2006, pp. 328.
- [6]. C. Romero, S. Ventura, M. Pechenizkiy, & S. J. d. R. Baker (Eds.), "Handbook of educational data mining, data mining and knowledge discovery series", Florida: Chapman & Hall/CRC, 2010.
- [7]. S. D'Mello, & A. Graesser, "Mining bodily patterns of affective experience during learning" In *Proceedings of the 3rd international conference on educational data mining*, 2010, pp. 31–40.
- [8]. R. S. J. D. Baker, S. M. Gowda, M. Wixon, J. Kalka, A. Z. Wagner, A. Salvi, V. Alevan, G. W. Kusbit, J. Ocumpaugh, & L. Rossi, "Towards sensor-free affect detection in cognitive tutor algebra", In *Proceedings of the 5th international conference on educational data mining*, 2012, pp. 126–133.
- [9]. A. Pena-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works", *Expert Systems with Application*, vol. 41, pp. 1432-1462, 2014.
- [10]. A.T. CORBETT, "Cognitive Computer Tutors: Solving the Two-Sigma Problem", In *Proceedings of the International Conference on User Modeling*, 2001, pp. 137-147.
- [11]. H. Guruler, A. Istanbulu, & M. Karahasan, "A new student performance analysing system using knowledge discovery in higher educational databases" *Computers & Education*, vol. 55, issue 1, pp. 247-254, 2010.
- [12]. J. D. Novak and A. J. Cañas, "The theory underlying concept maps and how to construct and use them," Inst. Human Mach. Cogn., Pensacola. FL, Tech. Rep. IHMC CmapTools, 2006-01, 2006.
- [13]. K. kundu & R. Punia, " Survey on educational data mining task", In *Proceeding of National Conference on Future Innovations in Computing Technologies & Machine Learning*, 2017, pp. 68-73.