ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
Available online at www.ijarbest.com

International Journal of Advanced Research in Biology Ecology Science and Technology (IJARBEST)
Vol. 2, Special Issue I, January 2016 in association with
KARUR VELALAR COLLEGE OF ARTS AND SCIENCE FOR WOMEN, KARUR
Conference on Emerging Trends and Functional Applications of Computer Technology - 12th January 2016

DATA MINING APPROACHES IN EDUCATION: A RESEARCH ANALYSIS

L. VIGNESWARAN[1]
Research Scholar (Part-Time),Bharathiyar University,Coimbatore,
TamilNadu, India.
vigneswaran.research@gmail.com

Dr. Mrs. N. NAGADEEPA[2]
Principal, Karur Velalar College of Arts and Science For Women, Karur, TamilNadu, India.
nagadeepa1012@gmail.com

Abstract — Educational Data Mining (EDM) is rising as a research area with a suite of computational and psychological methods and research approaches for understanding our student learn. Educational Data Mining(EDM) is concerned with developing methods to discover the unique types of data in educational settings and, using these methods, to better understand students and the settings in which they learn. Educational Data Mining (EDM) is used to study the data available in the educational field and bring out the hidden knowledge from it. Educational Data mining (EDM) methods often differ from methods from other data mining prose. This discipline focuses on analyzing Educational data to develop models for improving learning experiences and improving institutional efficiency. A literature review on educational data mining topics such as student maintenance and slow destruction, personal recommender systems within education, and how data mining can be used to analyze course management system data.

Keywords—Educational Data Mining, institutional efficiency, learning experience

1. INTRODUCTION

Two areas that are specific to the use of big data in education are educational data mining and learning analytics. Generally, educational data mining looks for new patterns in data and develops new algorithms and/or new models, while learning analytics applies known predictive models in instructional systems. Institutions are increasingly held accountable for success. One response to this process is finding new ways to apply analytical and data mining methods to educationally related data. The emerging field of educational data may examines the unique way of apply data mining methods to solve educational related problems[1]. Educational Data Mining (EDM) develops methods and applies techniques from statistics, machine learning, and data mining to analyze data collected during teaching and learning. EDM tests learning theories and informs educational practice. Learning analytics applies techniques from information science, sociology, psychology, statistics, machine learning, and data mining to analyze data collected during education administration and services, teaching, and learning. Learning analytics creates applications that directly influence educational practice.

In Educational data mining research, the following aspects are the goals for the research:

1. Predicting students' future learning behavior by creating student models that incorporate such detailed information as students' knowledge, motivation and attitudes.

2. Discovering or improving domain models that characterize the content to be learned and optimal instructional sequence.

3. Studying the effects of different kind of pedagogical support that can be provided by learning software.

4. Advancing scientific knowledge about learning and learners through building computational models that incorporate models of the student and the domain.[2]

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
Available online at www.ijarbest.com
International Journal of Advanced Research in Biology Ecology Science and Technology (IJARBEST)
Vol. 2, Special Issue I, January 2016 in association with
KARUR VELALAR COLLEGE OF ARTS AND SCIENCE FOR WOMEN, KARUR
Conference on Emerging Trends and Functional Applications of Computer Technology - 12th January 2016

There are a wide variety of current methods popular within educational data mining. These methods fall into the following general categories: prediction, clustering, relationship mining, discovery with models, and distillation of data for human judgment. The first three categories are largely acknowledged to be universal across types of data mining. The fourth and fifth categories achieve particular importance within educational data mining. [3].

2. EDUCATIONAL DATA

Decision-making in the field of academic planning involves extensive analysis of huge volumes of educational dat. Data's are generated from heterogeneous sources like diverse and varied uses in, diverse and distributed, structured and unstructured data. These data's are mostly generated from the offline or online source.

Online Data are generated from the geographically separated stake holder of the education, distance educations, web based education, computer supported collaborative learning, social networking sites and online group forum.

Offline Data are generated from traditional and modern classroom interaction interactive teaching/learning environments, learner/educators information, students attendance, emotional data, course information, data collected from the academic section of an institution etc.

Uncertain Data are generated from scientific measurement techniques and heterogeneity in designing data warehouse, sensor generated data, privacy preservation process data, summarization of data.

2.1 EDUCATIONAL TASK

It is a continual process for formation of Vision and Mission of an institution, to nurture the talent of students which addresses issues in a responsive, ethical and innovative manner to meet the academic and administrative objectives. This task can divide into two types:

Decision making task-active participation of the hybrid group of stakeholder to fulfill administrative oriented objectives

Learner based task- active participation of primary stakeholders to fulfill academic objectives.[7]

3. DATA MINING FUNCTIONALITIES IN EDM

To accomplish the EDM research goals, the data mining functionalities are playing the vital role in educational data mining research. The six categories of technical methods described below-

1. Classification- is the process of finding a model (or function) that describes and distinguishes data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose label is unknown. For example, in a student database, we would like to classify a large set of items, based on the three kinds of responses to a student academic: excellent, average, and tedious.[4]
2. Prediction- Classification predicts categorical(discrete, unordered) labels, prediction models continuous-valued functions. That is, it is used to predict missing or unavailable numerical data values rather than class labels. The term prediction may refer to both numeric prediction and class label prediction. Examples of using prediction include detecting such student behaviors as when they are gaming the system, engaging in off-task behavior, or failing to answer a question correctly despite having a skill.
3. Clustering- Clustering analyzes data object without consulting a known class label. The objects are clustered or grouped based on the principle of maximizing the intra class similarity and minimizing the interclass similarity. Clustering refers to finding data points that naturally group together and can be used to split a full dataset into categories. Examples of clustering applications are grouping students based on their learning difficulties and interaction patterns, such as how and how much they use tools in a learning management system, and grouping users for purposes of recommending actions and resources to similar users
4. Mining frequent relationships- involves discovering relationships between variables in a dataset and encoding

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
Available online at www.ijarbest.com
International Journal of Advanced Research in Biology Ecology Science and Technology (IJARBEST)
Vol. 2, Special Issue I, January 2016 in association with
KARUR VELALAR COLLEGE OF ARTS AND SCIENCE FOR WOMEN, KARUR
Conference on Emerging Trends and Functional Applications of Computer Technology - 12th January 2016

them as rules for later use. For example, relationship mining can identify the relationships among students borrowing books in Library.

A frequent itemset can be used for finding student mistakes that co-occur, associating content with user types to build recommendations for content that is likely to be interesting, or for making changes to teaching approaches. These techniques can be used to associate student activity, in a learning management system or discussion forums, with student grades or to investigate such questions as why students' use of practice tests decreases over a semester of study.

A frequent sequential pattern mining builds rules that capture the connections between occurrences of sequential events. For example, finding temporal sequences, such as student mistakes followed by help seeking. This could be used to detect events, such as students regressing to making errors in mechanics when they are writing with more complex and critical thinking techniques, and to analyze interactions in online discussion forums.

Key educational applications of mining frequent relationship include discovery of associations between student performance and course sequences and discovering which pedagogical strategies lead to more effective or robust learning. This is called teaching analytics—is of growing importance and is planned to help researchers build automated systems that model how effective teachers operate by mining their use of educational systems .

5. Refining for human judgment- is a technique that involves depicting data in a way that enables a human to quickly identify or classify features of the data. This area of educational data mining improves machine-learning models because humans can identify patterns in, or features of, student learning actions, student behaviors, or data involving collaboration among students.

6. Discovery with models- is a technique that involves using a validated model of a trend as a component in further analysis. For example, that categorized student activity from basic behavior data: students' interactions with a game-like learning environment that uses learning by teaching.

4. COURSE MANAGEMENT SYSTEMS IN EDM

A large number of researchers within EDM and how they can be improved to support student learning outcomes and student success. One research team developed a simplified data mining toolkit course management system and allows non experts users to get data mining information for their coursed.

A tool kit allows teachers to collaborate with each other and share results. This research is important because most data mining tools are complicated and require deep expertise in data mining tools, methods and processes, statistics, and machine learning algorithms. This studies follows a typical data mining process, thus it is quantitative. The then an application of specific data mining techniques, and then a post research and application contributions will allow non data mining activities. It is clear that additional mining tools more accessible to non-technical users.

Course management systems such as open source Moodle can be mined for usage data to find interesting patterns and trends in student online behaviour. The benefit to mining usage data is that it contains data about every user activity, such as testing, quizzes, reading, and discussion posts. Importance of pre-processing the data and then discuss specifics on how to apply data mining techniques to Moodle data. Their research results data, even if a reader does not have much experience in this area. The authors also use both Keel and Weka as their data mining software packages.

Data mining can be used in such a way as to customize learning activities for each individual student. Data mining was used through a course on English language instruction. Instead of having static course content, the course adapts to student learning, taking their through the course at their own pace. This was an effort for each student, and was a success. This research could be applied to other types of courses where students begin a course with varying levels of competency, e.g., a computer programming course.

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
Available online at www.ijarbest.com

International Journal of Advanced Research in Biology Ecology Science and Technology (IJARBEST)
Vol. 2, Special Issue I, January 2016 in association with
KARUR VELALAR COLLEGE OF ARTS AND SCIENCE FOR WOMEN, KARUR
Conference on Emerging Trends and Functional Applications of Computer Technology - 12th January 2016

5. EDM APPLICATION AREAS

Romero & Ventura [5] reviewed the EDM articles and suggested that future EDM research focus on the following aspects: -

1. Integrate EDM tools with e-learning systems
2. Standardize data and models.
3. Make EDM tools easier for educators and non-expert users.
4 Customize traditional mining algorithms for educational context. [6]

6. CONCLUSION

Educational data mining (EDM) is an area full of exciting opportunities for researchers and practitioners. This field assists higher educational institutions with efficient and effective ways to improve institutional effectiveness and student learning. Data mining is a significant tool for helping organizations enhance decision making and analyzing new patterns and relationships among a large amount of data. Working with big data using data mining and analytics is rapidly becoming common in the commercial sector. Tools and techniques once confined to research laboratories are being adopted by forward-looking industries, most notably those serving end users through online systems. Higher education institutions are applying learning analytics to improve the services they provide and to improve visible and measurable targets such as grades and retention.Future research can examine how widespread the adoption of educational data mining might be. Currently, it appears that research in this area is isolated and we do not know the exact extension of how institutions might be using data mining for enhancing student learning or improving related educational processes. It would be interesting to determine if there are barriers that prevent institutions from establishing EDM initiatives. There are a few case studies on how EDM is applied to admissions and enrollment, but further work needs to be done because those case studies seem isolated from the mainstream EDM work.

REFERENCES

[1] Richard A. Huebner, Norwich University, *Research in Higher Education Journal*.

[2] Marie Bienkowski, Mingyu Feng, Barbara Means, *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics* ,Center for Technology in Learning SRI International, Oct-2012.

[3] Ryan S.J.d. Baker, *Data Mining for Education*, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.

[4] Jiawei Han, Micheline Kamber, Jian Pei, *Data Mining Concepts and Techniques- Third Edition*.

[5] Baker, R. S., A. T. Corbett, K. R. Koedinger, and A. Z. Wagner. 2004. "Off-Task Behavior in the Cognitive Tutor Classroom: When Students 'Game the System.'" In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04). New York, NY: Association for Computing Machinery, 383–390.

[6] S.Lakshmi Praba, *EDUCATIONAL DATA MINING APPLICATIONS*, Operations Research and Applications: An International Journal (ORAJ), Vol. 1, No. 1, August 2014.

[7] Rajni Jindal, Malaya Dutta Borah- A Survey on Educational Data Mining and Research Trends, Department of Computer Engineering, Delhi Technological University, New Delhi, India, International Journal of Database Management Systems [IJDMS] Vol.5.No.3.June 2013.