Educational Data Mining (EDM): A Review

Mais Haj Qasem[1], Raneem Qaddoura[2], Bassam Hammo[1]

[1] King Abdullah II School for Information Technology, The University of Jordan, Amman, Jordan

[2] Philadelphia University, Amman, Jordan

E-mail: mais_hajqasem@hotmail.com, raneem.qaddoura@gmail.com, b.hammo@ju.edu.jo

Abstract

The application of various data mining (DM) tools and techniques in extracting useful information that are potentially valuable and significant is a trend in the research community that affects the education decision maker. Big data in education is also an important type of data to highlight because of the increasing number of online learning systems and open-source online courses, which makes "big data" important in this field. Educational DM (EDM) is useful in many different areas, including identifying at-risk students, identifying priority learning needs for different groups of students, increasing graduation rates, effectively assessing institutional performance, maximizing campus resources, and optimizing subject curriculum renewal. This study surveys the relevant studies in the EDM field.

Keywords: Data Mining, Big Data, Educational Data Mining

1. INTRODUCTION

Educational data mining (EDM) became an independent research area in 2008. It was established with the annual International Conference on Educational Data Mining and the Journal of Educational Data Mining. EDM is an area of scientific inquiry that focuses on the development of methods for use in discovering unique types of data from educational settings. These methods are also used to understand students and the learning settings better. EDM methods often differ from that of the broader data mining (DM) literature because they explicitly exploit the multiple levels of educational data hierarchy. Psychometrics literature methods are integrated with the machine learning and DM literatures for EDM use. How students choose educational software is an example of EDM. Simultaneously considering data at the keystroke, answer, session, student, classroom, and school levels may be worthwhile. Issues of time, sequence, and context also play important roles in the study of educational data.

Historical and operational data can be collected from educational institute databases. Student data can be personal or academic. Data can also be collected from e-learning systems, which are used by most institutes to store large amounts of information. The main objectives of higher education institutes are to provide quality education to their students and to improve the quality of managerial decisions. One way to achieve the highest level of quality in the higher education system is by extracting useful information from educational data to study the main attributes that may affect the performance of students. This knowledge can be used in providing helpful and constructive recommendations to academic planners in higher education institutes to enhance their decision-making process, improve the academic performance of students, trim down failure rate, understand the behaviour of students better, assist instructors, and improve teaching. EDM uses many techniques, such as decision tree, rule induction, neural networks, k-nearest neighbour, and naïve Bayesian. Knowledge in association rules, classifications, and clustering is discovered using these techniques.

The rest of this paper is organized as follows. The next section highlights the relationship between big data and EDM. Section 3 mentions the EDM process in general. Section 4 details the EDM methods that have been used in the field of education. Section 5 discusses the EDM application and uses. Section 6 highlights the related works in this field. The last section concludes this paper. students better, assist

2. BIG DATA

Online learning systems have increased in number. Tutors have been using these systems to manage courses, classes, and students. Meanwhile, students access and interact with the learning content anywhere and anytime using the Internet. With the advent of smartphones, the interaction of students with the learning content has increased. Traditional techniques are no longer suitable for "big data" because of the huge and complex amount of data generated.

Although many educational experience goals cannot be measured easily, data-intensive research and analysis in higher education can help in improving, controlling, and understanding the measurable goals. The breadth and depth of available data has the potential to fundamentally improve learning.

In our discussion, we highlight the use of massively open online courses (MOOCs) as a tool for data-intensive research and analysis in higher education. MOOCs illustrate that many types of big data can be collected in learning environments. Large amounts of data are gathered not only from many learners (broad between-learner data) but also on individual learning experiences (deep within-learner data). MOOC data includes longitudinal data (dozens of individual student courses over many years), rich social interactions (e.g., problem-solving videoconferences), and detailed data regarding specific activities (e.g., watching various segments of a video or individual actions in an educational game and problem solving). The depth of the data is determined not only by the amount of raw data on a learner but also by the availability of contextual information.

EDM offers a clear picture and better understanding of the learners and their learning processes. It uses DM techniques to analyse educational data and solve educational issues. The EDM extraction process is similar to that of other DM techniques, which extract interesting, interpretable, useful, and novel information from educational data. However, EDM is specifically concerned with developing methods to explore the unique types of data in educational settings [23]. Such methods are used to enhance the knowledge on educational phenomena, students, and their learning settings [24]. A large amount of data is available because of the increasing use of technology in educational systems. EDM provides a significant amount of relevant information [25], and the large volume of data gathered from distributed and diverse fields are used for decision-making and in the learning process in educational context.

Therefore, the following categories established by Romero et al. [26] are used as the main sources for EDM:

1) **Offline education**: It is also known as traditional education. In offline education, knowledge is transferred to learners on a face-to-face interaction basis. Data are collected using traditional methods, such as observation and questionnaires. It studies the cognitive skills of students and determines how they learn.

2) **E-learning and learning management systems (LMS):** These systems provide students with materials, instruction, communication, and reporting tools that allow them to learn by themselves. DM techniques can be applied to the data stored in the database systems.

Udupi et al. [27] identified and evaluated various e-learning models and the associated education technology paradigm. Their research further explored and proposed a new framework for big data integrations. They also discussed the scope of future research on DM and role of big data in e-learning environment. Furthermore, they explored the possibility embedding EDM features in the e-learning process and how the result and outcome can be used in various contexts. Their research also discussed how the DM concept helps create data clusters and how subsequent visualization and analysis can be done. Their result clearly identified how DM and big data integration in the e learning process can obtain advanced features, such as pattern generation, pattern analysis, predictive analysis, and knowledge discovery, which are necessary in identifying learning and learner’s needs in future e-learning revolutions.

3) **Intelligent tutoring systems (ITS) and adaptive educational hypermedia systems (AEHS):** These systems attempt to customize the student data based on student profiles. Thus, applying DM techniques is important in building user profiles. The data generated by these systems can be used in further research.

Amrieh et al. [28] proposed a new student’s performance prediction model based on DM techniques with new data attributes/features, which are called student’s behavioural features. These features are related to the learner’s interaction with the e-learning management system. The performance of the student’s predictive model is evaluated using a set of classifiers, namely, artificial neural networks, Naïve Bayesian, and decision tree. In addition, ensemble methods were applied to improve the performance of these classifiers. The common ensemble methods used in the literature were bagging, boosting, and random forest. The results revealed a strong relationship between the behaviours of learners and their academic achievement. The accuracy of their proposed model that uses behavioural features is better than those of ensemble methods where such features are removed.

3. EDM PROCESS

EDM uses an iterative, knowledge discovery process (see Fig-1). The process starts and takes place in educational environments, such as traditional classrooms, e-learning systems, and adaptive and intelligent Web-based educational systems. These educational environments generate raw data, including students' usage and interaction data, course information, and academic data. DM techniques, including clustering, classification, outlier, association, pattern matching, and text mining, which are discussed in the next section, are then applied to the raw data.

The EDM process consists of hypothesis formulation, testing, and refinement. Hypothesis is developed from various educational environments that create a large volume of data. The main process of EDM starts with data validation (i.e., finding the relationship between variables, parameters, and data items), which is also known as data pre-processing. After pre-processing using various DM techniques, tools are employed on the processed data, and results/interpretations are provided to various users of education. Further recommendations are suggested for addressing problems/task refinements.

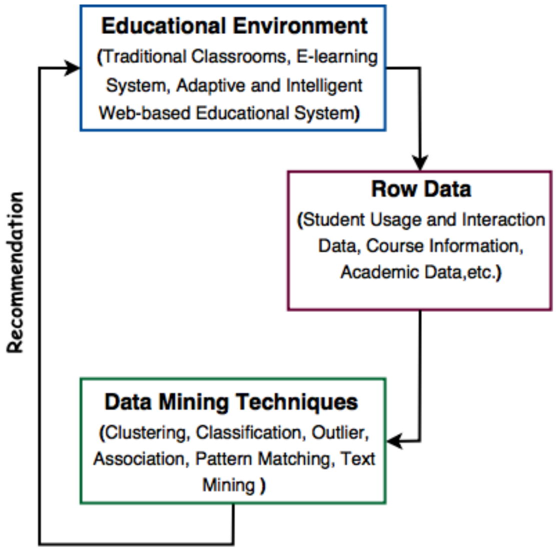

Fig. 1: EDM PROCESS

4. EDM METHODS

EDM methods are based on a variety of literature, including DM and machine learning, psychometrics, and other areas, such as statistics, information visualization, and computational modelling.

Romer Ventura [1] and Ryan Baker [2] categorized the works in EDM as prediction, clustering, and relationship mining.

4.1 Prediction

This technique is used in deriving a predicted variable (single variable) from predictor variables (combination of variables). Prediction is used in analysing student performance and drop out [3,4] and detecting student behaviour [5].

Sen et al. [55] developed models to predict secondary education placement tests using sensitivity analysis on predicting models (decision tree algorithm, support vector machine, neural network, and logistic regression). The following important predictors of placement test were identified: previous test experience, students with or without scholarship, number of siblings, and previous year's grade point average (GPA). The best predictors are decision tree analysis, followed by support vector machine and neural network. Logistic regression is the least suitable predictor.

Prediction is classified into three types:

4.1.1 Classification

Classification is a DM function that assigns items in a collection to target categories or classes. The goal of classification is to accurately predict the target class for each case in the data. For example, a classification model could be used to identify the credit risks of loan applicants as low, medium, or high. Some popular classification methods include logistic regression, support vector machines, and decision trees [6].

4.1.2 Regression

Regression is used to determine the relationship between the dependent variable (target field) and one or more independent variables. Dependent variables are values you want to predict, while independent variables are the basis of your prediction. Some popular regression methods within EDM include linear regression and neural networks [7].

4.1.3 Density Estimation

Probability density function is used to predict variables. Density estimator can be based on various kernel functions, including the Gaussian function.

Previous studies have shown how prediction techniques were used in EDM. Most of these studies used the classification techniques for analysing educational data.

Anuradha and Velmurugan [8] compared the performance of various classification techniques in terms of predicting the end of semester examination performance of students. A new system called fast boost decision tree algorithm was proposed to improve the accuracy of existing classification algorithms for the faster prediction of students based on their results. The proposed system is used to classify data and construct the decision tree by employing the top-down line search to test every attribute in the data set used for classification. The performance of the FBDT classifier was compared with C4.5 and 16 other classification algorithms.

Stephen [9] used EDM to predict the enrolment of students in science, technology, engineering, and mathematics (STEM) courses in higher educational institutions. The study examined the extent to which individual, sociodemographic, and school-level contextual factors help in pre-identifying successful and unsuccessful student enrolments to the STEM disciplines in the higher education institutions of Kenya. The Cross Industry Standard Process for Data Feature selection was used to rank the predictor variables using their importance for further analysis.

Kaur et al. [10] focused on identifying the slow learners among students and displaying the result through a predictive DM model using classification-based algorithms. A real-world data set from a high school was taken, and desired potential variables were filtered using an opensource tool called Waikato Environment for Knowledge Analysis (WEKA). The data set of student academic records was tested and applied on various classification algorithms, such as multilayer perception, Naïve Bayes, social media optimization, J48, and REP-Tree using WEKA. As a result, statistics were generated based on all classification algorithms, and all five classifiers were classified to predict the accuracy and to find the best-performing classification algorithm.

Devasia and Vinushree [11] proposed a system with a Web-based application that uses the Naive Bayesian mining technique in extracting useful information. Their experiment was conducted in Amrita Vishwa Vidyapeetham, Mysuru on 700 students with 19 attributes. Naive Bayesian algorithm provides more accurate results for comparison and prediction than other methods, such as regression, decision tree, and neural networks. The system aims at increasing the success graph of students using Naive Bayesian and maintains all student admission, course, subject, student marks, and attendance details.

4.2 Clustering

Clustering is an unsupervised classification process used for grouping similar objects into classes [12]. Data items are partitioned into groups or subsets (clusters) based on their locality and connectivity within N-dimensional space. In EDM, clustering has been used to group students according to their learning [13].

Macedo et al. [14] characterized the profiles of students based on an educational database. They assessed the application of K-means and C-means algorithms in clustering. The profiles were classified according to the Davies–Bouldin metrics and the gap statistic. The experiment results revealed that K-means is inefficient for the database. In contrast, C-means exhibited a satisfactory result, which was supported by Spearman's rank correlation coefficient. Educational platforms have been improving their functionalities to identify and understand several problems. Some problems have to be highlighted as how to understand the users and how to identify their needs. They showed that answering these questions for a confidential data set with more than 300 thousand students and 17 subjects is possible.

Rodrigues et al. [15] analysed the effectiveness of EDM techniques, specifically the cluster analysis, in identifying the engagement patterns of students in the MOOC course mode. Their analysis demonstrated the use of the hierarchical clustering method (Ward clustering) and the non-hierarchical clustering method (K-means) to analyse the engagement behaviour characteristics, which involved carrying out activities and interactions via the forum. The interaction patterns made in discussion murals at the Openredu platform and the data access and completeness activities were considered. The insights in this study are used by MOOC designers as indicators to satisfy the engagement pattern and design interface diversities that guide the design of adaptive strategies, which allow increasing engagement and fostering a better learning experience.

Ramos et al. [16] described the data analysis of Moodle's database of a beginner class at Distance Education of the Federal University using distinct EDM clustering methods. They carried out clustering using hierarchical and non-hierarchical methods in different groups of students, according to their interaction and performance characteristics. The analysis indicates that perceiving the groups obtained, perceiving the similarity between the results of the methods used, confirming the acquired knowledge from the clustering, and demonstrating that the choice of method has little influence on the knowledge obtained from interactions and the performance of students in the course are possible.

4.3 Relationship Mining

In relationship mining, the goal is to discover relationships between variables in a data set with a large number of variables. Relationship mining may attempt to determine which variables are most strongly associated with a single variable of particular interest or to discover which relationships between any two variables are the strongest.

Relationship mining is classified into four types:

4.3.1 Association Rule Mining

Association rule mining is a procedure that is meant to identify frequent patterns, correlations, associations, or causal structures from data sets found in various kinds of databases, such as relational databases, transactional databases, and other forms of data repositories, to determine the mistakes that students often make while solving exercises [17], [18].

4.3.2 Correlation Mining

This method is used to find linear correlations between variables (positive or negative). Correlation analysis is used to find the most strongly correlative attributes and discover interesting or unusual patterns of dependency among a large number of variables (sequences, signals, images, and videos).

4.3.3 Sequential Pattern Mining

This method is a DM topic of concern in finding statistically relevant patterns between data examples where the values are delivered in a sequence. The values are usually presumed discrete, and thus time series mining is closely related, but usually considered a different activity. Wang et al. proposed a four-phase learning portfolio mining approach [19].

4.3.4 Casual DM

This method is used to find causal relationships between variables by analysing the covariance of two events or by using the information on how one of the events was triggered.

Previous studies have shown how relationship mining were used to discover relationships between variables in a data set.

Bambrah et al. [20] analyzed the potential use of one of the DM techniques called association rule mining in enhancing the quality of student performances. The extracted rules help predict the performance of students and identify poor, good, and excellent students. The performance report of the student also helps improve the student result. Their system also helps identify students who need special attention to reduce the failure ratio and to apply appropriate action for the next semester's examination.

Kumar and Chadha [21] conducted a case study for a university that hopes to improve the quality of education by analysing the data and determining the factors that affect the academic results to increase the success chances of students. In this perspective, they used association rules discovery techniques. They also showed the importance of data pre-processing in data analysis, which has a significant impact on the accuracy of the predicted results. The mined association rules revealed various factors, such as student's interest, curriculum design, and teaching and assessment methodologies, that contributed to the failure of students to attain a satisfactory level of performance in the post graduation level.

Mahajan and Namdeo [22] presented a study that applied a predictive a priori algorithm in finding the hidden association between different subjects. For a student to successfully pass the MCA course, a sound knowledge of programming is a must. The associated model will help future students who are getting similar grades in elementary algorithm, procedural-oriented language, and object-oriented programming. In the curriculum design, two associated subjects should be placed in adjacent semesters so that the expertise gained in the previous subject can be used in the associated subject in the coming semester. The system will warn such students at a very early stage, and corrective precautionary measures can be taken. Teachers can also appropriate special attention where it is necessary. They suggested that in the curriculum design, two associated subjects should be placed in adjacent semesters.

5. EDM APPLICATION

EDM researchers study a variety of areas, including individual learning from educational software, computer supported collaborative learning, computer-adaptive testing (and testing more broadly), and the factors associated with student failure or non-retention in courses. Across these domains, one key area of application has been used in improving student models. Student models represent information about the characteristics or state of a student, such as the student's current knowledge, motivation, metacognition, and attitudes.

Modelling the individual differences of students in these areas enables software to respond to those individual differences, which significantly improves student learning [29]. EDM methods have enabled researchers to model a broader range of real-time, potentially relevant student attributes, including higher-level constructs. For instance, in recent years, researchers have used EDM methods to infer whether a student is playing the system [30], experiencing poor self-efficacy [31], off-task [32], or is bored or frustrated [33]. Researchers have also been able to extend student modelling beyond the educational software, toward figuring out what factors are predictive of student failure or no retention in college courses or in college altogether [4], [24], [34].

Discovering or improving models of a domain's knowledge structure is another EDM method application. By combining psychometric modelling frameworks with space-searching algorithms and the machine learning literature, a number of researchers have been able to develop automated approaches that discover accurate domain structure models directly from data. For instance, Barnes [35] developed an algorithm that can automatically discover a Q-Matrix from data. Desmarais and Pu [36] and Pavlik et al. [37] [38] developed algorithms for finding partial order knowledge structure models that explain the interrelationships of knowledge in a domain.

Some applications of EDM methods studied pedagogical support (both in learning software and in other domains, such as collaborative learning behaviours) toward discovering which types are most effective, in general or for different groups of students or for students in different situations [39], [40]. One popular method for studying pedagogical support is learning decomposition [39]. Learning decomposition fits exponential learning curves to performance data, thereby revealing that a student's future success is related to the amount and type of pedagogical support the student received up to that point. The relative weights for each type of pedagogical support in the best-fit model can be used to infer the relative effectiveness of each type of support to promote learning.

Lehr et al. [41] conducted one of the newest research that presented an EDM application for predicting undergraduate retention. The research provides valuable insight into data feature ranking, algorithm selections, and validation methods based on unique data types from educational settings. The data from a cohort of 972 students enrolled in 2008 at Embry–Riddle Aeronautical University (ERAU) were used to train and validate the predictive models. This research provides decision recommendations to ERAU and similar institutions to make timely interventions to improve retention.

6. RELATED LITERATURE

The computational overtures of EDM are used to analyse educational data to study edifying questions. Although many works exist to date, surveys like this provide a comprehensive resource of published papers on Scholastic Data Mining (EDM) from 2005 to 2015 that utilized a categorical DM technology.

Dutt et al. [44] advised researchers on how the puissance of massive didactic data of organizations can be utilized for strategic purposes using DM techniques. Meanwhile, it simplifies the system design, which takes cognizance from data, by utilizing sundry data mining approaches, such as clustering, classification, and prediction algorithms. They focused on consolidating the variants of clustering algorithms as applied in the EDM context.

Hannay et al. [45] utilized a pristine statistical method and concluded that efforts should be exerted on the elaboration of effects of personality on collaboration sundry measures, which may be habituated to prognosticate and influence performance.

Akinkunmi and Alo [47] conveyed DM methods to determine the performance of students in programming. The result of the study showed that a priori cognizance of physics and mathematics is essential for a student to excel in computer programming. This work is considerably useful in identifying students in early jeopardy, especially in profoundly and immensely colossal classes, and reminding the instructor to provide opportune advices in a timely manner.

Kumar [48] helped in learning and developing models for the growth of the education environment. A better understanding of student learning and the environment setting by decision makers in EDM was provided.

The relegation task to evaluate a student's performance was suggested by Baradwaj and Pal [50], and the decision tree method was utilized because of the many approaches used for data classification. They extracted knowledge that describe student performances in the end semester's examination. The study revealed that recognizing the dropouts and students who require special attention and requiring the edifier to provide congruous advising and counselling are important.

7. CONCLUSION

EDM has evolved as a multidisciplinary scientific learning area, which is rich in data, methods, tools, and techniques, that can be exploited to provide a better learning environment for users in the educational context. This survey helps in developing a good EDM method and is an overview tool, which is currently used to introduce improvements in teaching and predicting the academic performance of students.

REFERENCES

[1] ROMERO, C. and VENTURA, S. 2007. Educational Data Mining: A Survey from 1995 to 2005. Expert Systems with Applications 33, 125-146.

[2] Baker, R.S.J.D. in press.‖ Data Mining For Education‖, In International Encyclopedia of Education (3rd edition), B. MCGAW, PETERSON, P., BAKER Ed. Elsevier, Oxford, UK. 2009.

[3] P.V.Praveen Sundar.‖ A Comparative Study For Predicting Student's Academic Performance Using Bayesian Network Classifiers‖, IOSR Journal of Engineering (IOSRJEN) e-ISSN: 2250-3021, p-ISSN: 2278-8719 Vol. 3, Issue 2 (Feb. 2013), ||V1|| PP 37-42

[4] Dekker, G., Pechenizkiy, M., and Vleeshouwers J.‖Predicting Students Drop Out:A Case Study‖In Proceeding of the 2nd International Conference on Educational Data Mining,2009,pp.41-50

[5] Baker R.S.J.d, Gowda SM, Corbett AT. ―Automatically detecting a student's preparation for future learning: help use is key‖. In: Fourth International Conference on Educational Data Mining. Eindhoven, The Netherlands, 2011, 179–188.

[6] Behrouz Minaei-Bidgoli I , Deborah A. Kashy ', Gerd Kortemeyer and William F. Punch,‖Predicting Student Performance: An Application of Data Mining Methods With an Educational Web-Based System‖, 33'd ASEEIIEEE Frontiers in Education Conference, 2003, 0-7803-7961.

[7] Ahmad A. kardan , Hamid Sadeghi, ‖A Decision Support System for Course Offering in Online Higher Education System‖ Int. Journal of Computational Intelligence Systems Vol. 6 No.5 , Sep 2013, 928-942.

[8] Anuradha, C., & Velmurugan, T. Fast Boost Decision Tree Algorithm: A novel classifier for the assessment of student performance in Educational data.

[9] Stephen, K. W. (2016). Data Mining Model for Predicting Student Enrolment in STEM Courses in Higher Education Institutions.

[10] Kaur, P., Singh, M., & Josan, G. S. (2015). Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector. *Procedia Computer Science*, *57*, 500-508.

[11] Devasia, T., Vinushree, T. P., & Hegde, V. (2016, March). Prediction of students performance using Educational Data Mining. In *Data Mining and Advanced Computing (SAPIENCE), International Conference on* (pp. 91-95). IEEE.

[12] Jain, A. K., Murty, M. N., & Flynn, P. J., ―Data clustering: A Review,‖ ACM Computing Surveys, 31(3), 1999, (pp. 264–323).

[13] Amershi, S., Conati, C., ―Automatic Recognition of Learner Groups in Exploratory Learning Environments,‖ Proceedings of ITS 2006,8th International Conference on Intelligent Tutoring Systems, 2006.

[14] Macedo, M. G., & Bastos-Filho, C. J. Clustering Users Based on the Capacity to Solve Questions in an Educational Platform.

[15] Rodrigues, R. L., Ramos, J. L. C., Silva, J. C. S., & Gomes, A. S. (2016). Discovery engagement patterns MOOCs through cluster analysis. *IEEE Latin America Transactions*, *14*(9), 4129-4135.

[16] Ramos, J. L. C., e Silva, R. E. D., Silva, J. C. S., Rodrigues, R. L., & Gomes, A. S. (2016). A Comparative Study between Clustering Methods in Educational Data Mining. *IEEE Latin America Transactions*, *14*(8), 3755-3761.

[17] Dai Shangping , Zhang Ping.‖ A Data Mining Algorithm In Distance Learning‖. IEEE, 2008, 978 -1-4244-1651.

[18] Merceron, A., & Yacef, K., ―Mining student data captured from a web-based tutoring tool: Initial exploration and results,‖ Journal of Inter

[19] Wang, W., Weng, J., Su, J., & Tseng, S., ―Learning portfolio analysis and mining in SCORM compliant environment,‖ In ASEE/ IEEE frontiers in education conference, 2004, (pp. 17–24).

[20] Bambrah, C., Bhandari, M., Maniar, N., & Munde, V. (2014). Mining Association Rules in Student Assessment Data. *International Journal of Advanced Research in Computer and Communication Engineering*, *3*(3).

[21] Kumar, V., & Chadha, A. (2012). Mining association rules in student's assessment data. *International Journal of Computer Science Issues*, *9*(5), 211-216.

[22] Mahajan, N., & Namdeo, J. Analysis of Student Result Data Using Association Rule Mining.

[23] S. K. Mohamad and Z. Tasir, "Educational data mining: A review," Procedia-Social and Behavioral Sciences, vol. 97, pp. 320–324, 2013.

[24] Baker, R. S. J. D. (2010). Data mining for education. *International encyclopedia of education*, *7*, 112-118.

[25] J. Mostow and J. Beck, "Some useful tactics to modify, map and mine data from intelligent tutors," Natural Language Engineering, vol. 12, no. 02, pp. 195–208, 2006.

[26] C. Romero, S. Ventura, P. G. Espejo, and C. Hervas, "Data mining ´ algorithms to classify students.," in EDM, pp. 8–17, 2008.

[27] Udupi, P. K., Sharma, N., & Jha, S. K. (2016, September). Educational data mining and big data framework for e-learning environment. In *Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), 2016 5th International Conference on* (pp. 258-261). IEEE.

[28] Amrieh, E. A., Hamtini, T., & Aljarah, I. (2016). Mining Educational Data to Predict Student's academic Performance using Ensemble Methods. *International Journal of Database Theory and Application*, *9*(8), 119-136.

[29] CORBETT, A.T. 2001. Cognitive Computer Tutors: Solving the Two-Sigma Problem. In Proceedings of the International Conference on User Modeling, 137-147

[30] BAKER, R.S.J.D. in press. Data Mining For Education. In International Encyclopedia of Education (3rd edition), B. MCGAW, PETERSON, P., BAKER Ed. Elsevier, Oxford, UK.

[31] MCQUIGGAN, S., MOTT, B. and LESTER, J. 2008. Modeling Self-Efficacy in Intelligent Tutoring Systems: An Inductive Approach. User Modeling and User-Adapted Interaction 18, 81-123

[32] BAKER, R.S.J.D. 2007. Modeling and Understanding Students' Off-Task Behavior in Intelligent Tutoring Systems. In Proceedings of the ACM CHI 2007: Computer-Human Interaction conference, 1059-1068.

[33] D'MELLO, S.K., CRAIG, S.D., WITHERSPOON, A.W., MCDANIEL, B.T. and GRAESSER, A.C. 2008. Automatic Detection of Learner's Affect from Conversational Cues. User Modeling and User-Adapted Interaction 18, 45-80.

[34] SUPERBY, J.F., VANDAMME, J.-P. and MESKENS, N. 2006. Determination of factors influencing the achievement of the first-year university students using data mining methods. In Proceedings of the Workshop on Educational Data Mining at the 8th International Conference on Intelligent Tutoring Systems (ITS 2006), 37-44.

[35] BARNES, T. 2005. The q-matrix method: Mining student response data for knowledge. In Proceedings of the AAAI-2005 Workshop on Educational Data Mining.

[36] BARNES, T., DESMARAIS, M., ROMERO, C. and VENTURA, S. 2009. Educational Data Mining 2009: 2nd International Conference on Educational Data Mining, Proceedings, Cordoba, Spain.

[37] PAVLIK, P., CEN, H. and KOEDINGER, K.R. 2009. Learning Factors Transfer Analysis: Using Learning Curve Analysis to Automatically Generate Domain Models. In Proceedings of the 2nd International Conference on Educational Data Mining, 121-130.

[38] PAVLIK, P., CEN, H., WU, L. and KOEDINGER, K. 2008. Using Item-type Performance Covariance to Improve the Skill Model of an Existing Tutor. In Proceedings of the 1st International Conference on Educational Data Mining, 77-86.

[39] BECK, J.E. and MOSTOW, J. 2008. How who should practice: Using learning decomposition to evaluate the efficacy of different types of practice for different types of students. In Proceedings of the 9th International Conference on Intelligent Tutoring Systems, 353-362.

[40] PECHENIZKIY, M., CALDERS, T., VASILYEVA, E. and DE BRA, P. 2008. Mining the Student Assessment Data: Lessons Drawn from a Small Scale Case Study. In Proceedings of the 1st International Conference on Educational Data Mining, 187-191.

[41] Lehr, S., Liu, H., Kinglesmith, S., Konyha, A., Robaszewska, N., & Medinilla, J. (2016, December). Use Educational Data Mining to Predict Undergraduate Retention. In *Advanced Learning Technologies (ICALT), 2016 IEEE 16th International Conference on* (pp. 428-430). IEEE.

[42] Cristobel Romero, Sebastian 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, no. 6, november 2010

[43] 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

[44] Ashish Dutt, Saeed Aghabozrgi, Maizatul Akmal Binti I smail, and Hamidreza Mahr oeian, "Clustering Algorithms Applied in Educational Data Mining", International Journal of Information and Electronics Engineering, Vol. 5, No. 2, March 2015

[45] Jo E. Hannay, Erik Arisholm,Harald Engvik, and Dag I.K. Sjøberg, "Effects of Personality on Pair Programming", IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 36, NO. 1, JANUARY/FEBRUARY 2010

[46] Agathe MERCERON, and Kalina YACEF, "Educational Data Mining: a Case Study ", Proceedings of the 12th international Conference on Artificial Intelligence in Education AIED 2005.

[47] 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

[48] Jasvinder Kumar, A Comprehensive Study of Educational Data Mining, International Journal of Electrical Electronics & Computer Science Engineering Special Issue - TeLMISR 2015, ISSN: 2348-2273

[49] Rajni Jindal and Malaya Dutta Borah, "A Survey On Educational Data Mining And Research Trends",International Journal of Database Management Systems (IJDMS) Vol.5, No.3, June 2013

[50] Brijesh Kumar Baradwaj, Saurabh Pal, "Mining Educational Data to Analyze Students Performance", International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011

[51] M.Jayakameswaraiah, S.Ramakrishna, "A Study on Prediction Performance of Some Data Mining Algorithms", International Journal of Advance Research in Computer Science and Management Studies, ISSN: 2321-7782 (Online) Volume 2, Issue 10, October 2014

[52] Suchita Borkar, K. Rajeswari, "Predicting Students Academic Performance Using Education Data Mining ", International Journal of Computer Science and Mobile Computing, ISSN 2320–088X IJCSMC, Vol. 2, pg.273– 279 Issue. 7, July 2013

[53] A.Dinesh Kumar ,Dr.V.Radhika, "A Survey on Predicting Student Performance", International Journal of Computer Science and Information Technologies, Vol. 5 (5) , 2014,ISSN:0975-9646, 6147- 6149

[54] Azwa Abdul Aziz, Nor Hafieza Ismail and Fadhilah Ahmad, "First Semester Computer Science Students' Academic Performances Analysis by Using Data Mining Classification Algorithms", International Conference on Artificial Intelligence and Computer Science (AICS 2014), September 2014

[55] Sen, A., & Srivastava, M. (2012). Regression analysis: theory, methods, and applications. Springer Science & Business Media.