Review Paper on Educational Data Mining

Mr. Pradeep Nayak[1], Mohammed Sufiyan[2], Mohan Raju. V [3], Monisha. N. S.[4], Moollya Gautami Bhaskar [5]

Assistant Professor, Department of Information Science and Engineering[1]

Students, Department of Information Science and Engineering[2,3,4,5]

Alva's Institute of Engineering and Technology, Mijar, Mangalore, Karnataka, India

Abstract: *Education and computer science are both involved in the burgeoning inter-disciplinary research field known as Educational Data Mining (EDM). EDM uses data mining software and ways to extract meaningful and practical data from big educational databases. EDM introduces better and more efficient learning techniques in an effort to enhance educational processes. The term "EDM methods" refers to a group of techniques for creating models and applications. This page provides a thorough literature review on EDM techniques. The essay also covers EDM research problems and trends.This EDM insight aims to provide researchers interested in furthering the field of EDM with useful and valuable information.*

Keywords: Educational Data Mining (EDM); Prediction; Relationship Mining; Structure Discovery are some keywords

REFERENCES

[1]. J. Han, J. Pei, and M. Kamber, Data mining: concepts and techniques. Elsevier, New York, USA, 2011.

[2]. Website of International Educational Data Mining Society (last accessed on 19 February 2020). Available at www.educationaldatamining.org.

[3]. P. Nithya, B. Umamaheswari, and A. Umadevi, "A survey on educational data mining in field of education," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 5, no. 1, 2016.

[4]. R. S. Baker and P. S. Inventado, "Educational data mining and learning analytics," in Learning Analytics, pp. 61–75, Springer, New York, USA, 2014.

[5]. C. Romero and S. Ventura, "Educational data mining: A survey from 1995 to 2005," Expert Systems with Applications, vol. 33, no. 1, pp. 135– 146, 2007.

[6]. C. Romero and 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, no. 6, pp. 601–618, 2010.

[7]. R. S. Baker and K. Yacef, "The state of educational data mining in 2009: A review and future visions," JEDM-Journal of Educational Data Mining, vol. 1, no. 1, pp. 3–17, 2009.

[8]. R. S. Baker, "Data mining for education," International encyclopedia of education, vol. 7, no. 3, pp. 112–118, 2010

[9]. G. Kashyap and E. Chauhan, "Review on educational data mining tech- niques," International Journal of Advance Technology in Engineering and Science, vol. 3, no. 11, 2015.

[10]. C. Romero and S. Ventura, "Data mining in education," Wiley Interdisci- plinary Reviews: Data Mining and Knowledge Discovery, vol. 3, no. 1, pp. 12–27, 2013.

[11]. C. Vieira, P. Parsons, and V. Byrd, "Visual learning analytics of educational data: A systematic literature review and research agenda," Computers & Education, vol. 122, pp. 119–135, 2018.

[12]. A. Shukla, "Ph.d. thesis report - a study of relationship between symbols and sentiments for management of annotated academic resources," tech. rep., MNNIT Allahabad, India, 2014.

[13]. A. Merceron and K. Yacef, "Educational data mining: a case study," in AIED, pp. 467–474, 2005.

[14]. C. Romero, S. Ventura, and E. Garc´ıa, "Data mining in course manage- ment systems: Moodle case study and tutorial," Computers & Education, vol. 51, no. 1, pp. 368–384, 2008

[15]. R. Paiva, I. I. Bittencourt, W. Lemos, A. Vinicius, and D. Dermeval, "Visualizing learning analytics and educational data mining outputs," in International Conference on Artificial Intelligence in Education, pp. 251–256, Springer, 2018.

[16]. B. Minaei-Bidgoli, D. A. Kashy, G. Kortemeyer, and W. F. Punch, "Predicting student performance: an application of data mining methods with an educational web-based system," in Frontiers in Education, 2003. FIE 2003 33rd Annual, vol. 1, pp. T2A–13, IEEE, 2003.

[17]. B. K. Baradwaj and S. Pal, "Mining educational data to analyze students' performance," Computing Research Repository - CoRR, vol. abs/1201.3417, 2012.

[18]. C. Romero, S. Ventura, P. G. Espejo, and C. Hervás, "Data mining algorithms to classify students," in Educational Data Mining, 2008.

[19]. P. Kaur, M. Singh, and G. S. Josan, "Classification and prediction based data mining algorithms to predict slow learners in education sector," Procedia Computer Science, vol. 57, pp. 500–508, 2015.

[20]. C. Romero, P. G. Espejo, A. Zafra, J. R. Romero, and S. Ventura, "Web usage mining for predicting final marks of students that use moodle courses," Computer Applications in Engineering Education, vol. 21, no. 1, pp. 135–146, 2013.

[21]. C. Márquez-Vera, A. Cano, C. Romero, and S. Ventura, "Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data," Applied Intelligence, vol. 38, no. 3, pp. 315–330, 2013.

[22]. C. Márquez-Vera, A. Cano, C. Romero, AY.M. Noaman, H. Mousa Fardoun, and S. Ventura, "Early dropout prediction using data mining: a case study with high school students," Expert Systems, vol. 33, no. 1, pp. 107–124, 2016.

[23]. E. B. Costa, B. Fonseca, M. A. Santana, F. F. de Araújo, and J. Rego, "Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory program- ming courses," Computers in Human Behavior, vol. 73, pp. 247–256, 2017.

[24]. M. I. Lopez, J. Luna, C. Romero, and S. Ventura, "Classification via clustering for predicting final marks based on student participation in forums," International Educational Data Mining Society, 2012

[25]. J. L. Olmo, C. Romero, E. Gibaja, and S. Ventura, "Improving metalearning for algorithm selection by using multi-label classification: A case of study with educational data sets," International Journal of Computational Intelligence Systems, vol. 8, no. 6, pp. 1144–1164, 2015.

[26]. A. Zafra and S. Ventura, "Multi-instance genetic programming for predicting student performance in web based educational environments," Applied Soft Computing, vol. 12, no. 8, pp. 2693– 2706, 2012.

[27]. D. A. Al Hammadi and M. S. Aksoy, "Data mining in education-an experimental study," International Journal of Computer Applications, vol. 62, Jan. 2013.

[28]. N. Delavari and S. Phon-Amnuaisuk, "Data mining application in higher learning institutions," Informatics in Education, 2008.

[29]. M. A. Sethi and M. C. Singh, "Data mining for prediction and classifica- tion of engineering students achievements using improved naïve bayes," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 6, pp. 966– 971, July 2017.

[30]. J. Zimmermann, K. H. Brodersen, H. R. Heinimann, and J. M. Buhmann, "A model-based approach to predicting graduate-leve performance using indicators of undergraduate-level performance," Journal of Educational Data Mining, vol. 7, no. 3, pp. 151–176, 2015.

[31]. P. Strecht, L. Cruz, C. Soares, J. Mendes-Moreira, and R. Abreu, "A comparative study of classification and regression algorithms for modelling students' academic performance.," International Educational Data Mining Society, 2015.

[32]. C. R. Beal, L. Qu, and H. Lee, "Classifying learner engagement through integration of multiple data sources," in Proceedings of the National Conference on Artificial Intelligence, vol. 21, pp. 151–156, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[33]. J. Ocumpaugh, R. Baker, S. Gowda, N. Heffernan, and C. Heffernan, "Population validity for educational data mining models: A case study in affect detection," British Journal of Educational Technology, vol. 45, no. 3, pp. 487–501, 2014.

[34]. A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," User Modeling and UserAdapted Interaction, vol. 4, pp. 253–278, Dec 1994.

[35]. A. Aleem and M. M. Gore, "C-BEAM: A confidence-based evaluation of MCQs for providing feedback to instructors," Computer Applications in Engineering Education, vol. 27, no. 1, pp. 112–127, 2019.

[36]. W. Tissera, R. Athauda, and H. Fernando, "Discovery of strongly related subjects in the undergraduate syllabi using data mining," in International Conference on Information and Automation, 2006. ICIA 2006., pp. 57– 62, IEEE, Shandong, China, 2006.

[37]. A. Buldu and K. U¨ c¸gu¨n, "Data mining application on students' data," Procedia-Social and Behavioral Sciences, vol. 2, no. 2, pp. 5251–5259, 2010.

[38]. E. Chandra and K. Nandhini, "Knowledge mining from student data," European Journal of Scientific Research, vol. 47, no. 1, pp. 156–163, 2010.

[39]. M. Chalaris, S. Gritzalis, M. Maragoudakis, C. Sgouropoulou, and A. Tsolakidis, "Improving quality of educational processes providing new knowledge using data mining techniques," Procedia-Social and Behavioral Sciences, vol. 147, pp. 390–397, 2014.

[40]. F. Shi, Q. Miao, and D. Mei, "The application of data association mining technology in university curriculum management," in 2012 IEEE Symposium on Robotics and Applications (ISRA), pp. 521– 524, IEEE, Kuala Lumpur, Malaysia, June 2012.

[41]. P. J. DeCarlo and N. Rizk, "The design and development of an expert system prototype for enhancing exam quality," International Journal of Advanced Corporate Learning (iJAC), vol. 3, pp. 10–13, 2010.

[42]. J. M. Luna, C. Romero, J. R. Romero, and S. Ventura, "An evolutionary algorithm for the discovery of rare class association rules in learning management systems," Applied Intelligence, vol. 42, no. 3, pp. 501–513, 2015.

[43]. C. Romero, A. Zafra, J. M. Luna, and S. Ventura, "Association rule mining using genetic programming to provide feedback to instructors from multiple-choice quiz data," Expert Systems, vol. 30, no. 2, pp. 162–172, 2013.

[44]. G. Dimić, B. Predić, D. Rančić, V. Petrović, N. Maček, and P. Spalević, "Association analysis of moodle e-tests in blended learning educational environment," Computer Applications in Engineering Education, vol. 26, no. 3, pp. 417–430, 2018

[45]. A. Aleem, A. Kumar, and M. M. Gore, "A study of manuscripts evolution to perfection," in 2nd International Conference on Adavanced Computing and Software Engineering - ICACSE 2019, pp. 278–282, 2019.

[46]. S.Y. Deng and X. Que, "Research on the teaching assessment of students of science and engineering teachers in a university," Computer Applications in Engineering Education, vol. 27, no. 1, pp. 5–12, 2019.

[47]. A. Aleem and M. M. Gore, "The choice is yours: The effects of optional questions in engineering examinations," Computer Applications in Engineering Education, vol. 27, no. 5, pp. 1087–1102, 2019.

[48]. P. Fournier-Viger, J. C.-W. Lin, R. U. Kiran, Y. S. Koh, and R. Thomas, "A survey of sequential pattern mining," Data Science and Pattern Recognition, vol. 1, no. 1, pp. 54–77, 2017.

[49]. J. Li, T. D. Le, L. Liu, J. Liu, Z. Jin, and B. Sun, "Mining causal association rules," in 13th IEEE International Conference on DataMining Workshops (ICDMW), pp. 114–123, IEEE, Dallas, Texas, USA, 2013.

[50]. S. Mani and G. F. Cooper, "A simulation study of three related causal data mining algorithms," in Eighth International Workshop on Artificial Intelligence and Statistics (AISTATS), pp. 73–80, Society for Artificial Intelligence and Statistics, Key West, Florida, USA, Jan. 2001.

[51]. A. H. Cairns, B. Gueni, M. Fhima, A. Cairns, and S. N. K. David, "Process mining in the education domain," International Journal on Advances in Intelligent Systems, vol. 8, no. 1 & 2, pp. 219–232, 2015.

[52]. A. Bogarín, R. Cerezo, and C. Romero, "A survey on educational process mining," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 8, no. 1, p. e1230, 2018.

[53]. T. S. Uhlmann, E. A. P. Santos, and L. A. Mendes, "Process mining applied to player interaction and decision taking analysis in educational remote games," in International Conference on Remote Engineering and Virtual Instrumentation, pp. 425–434, Springer, 2018.

[54]. A. Aleem, R. Srivastava, A. K. Singh, and M. M. Gore, "GCLOD: A clustering algorithm for improved intra-cluster similarity and efficient local outliers detection.," in DMIN, pp. 524–530, 2009.

[55]. R. Ahuja, A. Jha, R. Maurya, and R. Srivastava, "Analysis of educational data mining," in Harmony Search and Nature Inspired Optimization Algorithms, pp. 897–907, Springer, 2019.

[56]. A. Dutt, M. A. Ismail, and T. Herawan, "A systematic review on educational data mining," IEEE Access, 2017.

[57]. S. H. Ganesh and A. J. Christy, "Applications of educational data mining: A survey," in Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015 International Conference on, pp. 1–6, IEEE, 2015.

[58]. L. Ramanathan, G. Parthasarathy, K. Vijayakumar, L. Lakshmanan, and S. Ramani, "Cluster-based distributed architecture for prediction of student's performance in higher education," Cluster Computing, pp. 1–16, 2018.

[59]. M. Goyal and R. Vohra, "Applications of data mining in higher educa- tion," International Journal of Computer Science, vol. 9, no. 2, pp. 113–120, 2012.

[60]. M. C. Desmarais, "Mapping question items to skills with nonnegative matrix factorization," ACM SIGKDD Explorations Newsletter, vol. 13, no. 2, pp. 30–36, 2012.

[61]. M. I. Al-Twijri and A. Y. Noaman, "A new data mining model adopted for higher institutions," Procedia Computer Science, vol. 65, pp. 836–844, 2015.

[62]. H. Jeong and G. Biswas, "Mining student behavior models in learning- by-teaching environments," in Educational Data Mining, 2008.

[63]. G. Kostopoulos, S. Kotsiantis, and P. Pintelas, "Predicting student performance in distance higher education using semi-supervised techniques," in Model and Data Engineering, pp. 259–270, Springer, 2015.

[64]. G. Kostopoulos, A.-D. Lipitakis, S. Kotsiantis, and G. Gravvanis, "Pre- dicting student performance in distance higher education using active learning," in International Conference on Engineering Applications of Neural Networks, pp. 75–86, Springer, 2017.

[65]. R. Jindal and M. D. Borah, "A survey on educational data mining and research trends," International Journal of Database Management Systems (IJDMS), vol. 5, no. 3, p. 53, 2013.

[66]. C. Romero, J. R. Romero, and S. Ventura, "A survey on preprocessing education