

Educational Data Mining- A Review

¹Gaurav Kumar, ²Ashok Sharma

¹Research Scholar, Lovely Professional University, Punjab

²Associate Professor, Lovely Professional University, Punjab

Abstract — Data mining happens to be very helpful in understanding the behaviour of online learners. The data mining has the potential of extracting the hidden pattern of the information and analysing the same. This paper looks into the research conducted in the area of data mining and several limitations existing in the said research.

Keywords — Data Mining, E-learning

I. Introduction

Data mining is popularly known as knowledge discovery in database (KDD) and it is known for extracting hidden information from a large volume of data [1]. Data mining has shown its applications in the area of bioinformatics and e-commerce and also in the educational system which is known as educational data mining [2]. Educational data mining is defined by educational data mining community and EDM deals in student models [3].

Various research papers in the area of data mining in educational research have been collected [3,4,5,6]. The first review is related to the application of data mining techniques in various educational system from the year 1995 to 2005[5]. Another review was related to the e-learning problems [4].

II. Literature Review

Table 1: Review of educational data mining research from 2004 to 2012

Reference	Objective to achieve	Platform	Task	Source of Research	Number of citations
[7]	Extracting hidden pattern	TRAC	Method of Sequential Pattern	Conference proceedings	47
[8]	Automated Chat Analysis	Chat Platform	Classification Modelling Technique	Conference proceedings	24
[9]	Extracting learner preference in online	E-learning	Predictive Analytics	Conference proceedings	25

	learning				
[10]	Extracting the learners' assessment data	Online learning	Techniques like classification and clustering	Conference proceedings	10
[11]	Observing learners' online	Live online video streaming Platform	Techniques like Clustering	Journal issue	2
[12]	understanding disengagement problems	Online learning system	Techniques like Clustering	Journal issue	3
[13]	intelligent report system	Online learning	Techniques like Classification	Journal issue	1
[14]	Extracting learners data for similar groups	Digital Community System	Techniques like Clustering	Conference proceedings	97
[15]	Clustering and sequential pattern mining	TRAC	Techniques like clustering and sequential papattern	Conference proceedings	71

III. Methodology

The Table 1 shows that most of the research papers in educational data mining have used clustering [10,11,12,14,15]. It has also showed that many research papers are based on classification techniques [8,10,13]. Few research studies have used sequential pattern [7,15]. Prediction has also been used in many studies [9]. Many research papers have also used association rule analysis [10].

IV. Conclusion

It has been observed that most of the research has been done on the problem related to e-learning. The need of the hour is to perform research on collaboration learning by using platforms like facebook and other social networking platforms which are quite popular in students. The log data from these platforms and blogs can be analyzed by using google analytics [20,21].

REFERENCES

- [1] Witten, I.H. and Frank, E. 1999. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kauffman, San Francisco, CA.
- [2] Baker, R.S.J.d.: Data Mining for Education. In: McGaw, B., Peterson, P., Baker, E. (eds.) To appear in *International Encyclopedia of Education*, 3rd edn. Elsevier, Oxford (2010)
- [3] Baker, R. S. J. D., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3-17.
- [4] Castro, F., Vellido, A., Nebot, À., & Mugica, F. (2007). Applying data mining techniques to e-learning problems. In *Evolution of teaching and learning paradigms in intelligent environment* (pp. 183-221). Springer Berlin Heidelberg.
- [5] Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33(1), 135-146.
- [6] Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40(6), 601-618.
- [7] Kay, J., Maisonneuve, N., Yacef, K., & Zaïane, O. (2006) *Proceedings of the Workshop on Educational Data Mining at the 8th International Conference on Intelligent Tutoring Systems (ITS 2006)* (pp. 45-52)
- [8] Anjewierden, A., Kolloffel, B., & Hulshof, C. (2007). Towards educational data mining: Using data mining methods for automated chat analysis to understand and support inquiry learning processes. In *International Workshop on Applying Data Mining in e-Learning (ADML 2007)*.
- [9] Carmona, C., Castillo, G., & Millán, E. (2007). Discovering student preferences in e-learning. In *Proceedings of the international workshop on applying data mining in e-learning* (pp. 23-33).
- [10] Pechenizkiy, M., Calders, T., Vasilyeva, E., & De Bra, P. (2008). Mining the student assessment data: Lessons drawn from a small scale case study. *Educational Data Mining 2008*, 187.
- [11] He, W.(2012). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior*.

- [12] Shih, Y. C., Huang, P. R., Hsu, Y. C., & Chen, S. Y. (2012). A complete understanding of disorientation problems in web-based learning. *The Turkish Online Journal of Educational Technology*, 11(3).
- [13] Blagojevic, M., Micic, Z. (2012). A web-based intelligent report e-learning system using data mining techniques. *Computers & Electrical Engineering*.
- [14] Talavera, L., & Gaudioso, E. (2004). Mining student data to characterize similar behavior groups in unstructured collaboration spaces. In *Proceedings of the Artificial Intelligence in Computer Supported Collaborative Learning Workshop* at the ECAI 2004.
- [15] Perera, D., Kay, J., Koprinska, I., Yacef, K., & Zaïane, O. R. (2009). Clustering and sequential pattern mining of online collaborative learning data. *Knowledge and Data Engineering, IEEE Transactions on*, 21(6), 759-772.
- [16] Merceron, A., & Yacef, K. (2007). Revisiting interestingness of strong symmetric association rules in educational data. In *Proc. of Int. Workshop on Applying Data Mining in e-Learning, Crete, Greece* (pp. 3-12).
- [17] García, E., Romero, C., Ventura, S., & Calders, T. (2007, September). Drawbacks and solutions of applying association rule mining in learning management systems. In *Proceedings of the International Workshop on Applying Data Mining in e-Learning (ADML 2007)*, Crete, Greece (pp. 13-22).
- [18] Ingram, A. L. (2000). Using web server logs in evaluating instructional web sites. *Journal of Educational Technology Systems*, 28(2), 137-158.
- [19] Smith, B. L., & MacGregor, J. T. (1992). What Is Collaborative Learning?. National Center on Postsecondary Teaching, Learning, and Assessment at Pennsylvania State University
- [20] Churchill, D. (2009). Educational applications of Web 2.0: Using blogs to support teaching and learning. *British Journal of Educational Technology*, 40(1), 179-183.
- [21] DiVall, M. V., & Kirwin, J. L. (2012). Using Facebook to facilitate course-related discussion between students and faculty members. *American Journal of Pharmaceutical Education*, 76(2).