Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

An Intelligent Learning Analytics Model for Early Identification of At-Risk Students in Online and Blended Learning

Mohammed Abdul Khaled[1]*, Muhammad Nadeem[2], Ahmed Gamal Elsayed Hamed Ali Mansour[3], Ahmed Ali Mohammed Kulfod[4], Mohammed Abdul Aziz[5]
[1]*Osmania University, Hyderabad, India (khaledharvadx@gmail.com)*
[2]*Bahauddin Zakariya University, Multan, Pakistan (nomi647@hotmail.com)*
[3]*Al-Azhar University, Nasr City, Cairo, Egypt (ahmedshmes21@gmail.com)*
[4]*University of Science and Technology, Republic of Yemen (ahmed.kulfod@gmail.com)*
[5]*Jawaharlal Nehru Technological University Hyderabad, India (maaziz.mails@gmail.com)*
**Corresponding Author*

ARTICLE INFO

Received: 01 Oct 2022

Accepted: 20 Dec 2022

ABSTRACT

The ability to detect high-risk students who are likely to fail academically or drop out is an urgency in online and blended course set-ups. This research paper is an analytical exploration of early risk identification of learning analytics based on the use of secondary data, and without the creation of a new predictive model or a learning analytics system. Data published and anonymised about Coursera, edX and The Open University are studied to investigate how well academic risk can be detected timely with the help of established AI-assisted learning analytics processes. The analysis is performed in a systematic analytical process, which includes data preprocessing, engagement and assessment indicator construction, analysis of temporal behaviours, at-disengagement pattern detection, and identification of at-risk learning events. Learning analytics literature has suggested analytical procedures that are implemented at an analytical level. Back-testing of potential risk early detection is carried out through historical history back-testing by using known academic performance. The findings suggest that the prevalent factors of academic risk are sustained disengagement and decreasing learning activity. An at-risk learning episode is identified in 312 times out of which 84.6% are identified before the academic outcome is certain. Mean learning periods between early-risk signals and course failure / non-completion are on average 2.6, indicating that it is possible to obtain meaningful lead time on academic intervention. The level of risk also indicates that high-risk and severe-risk learning conditions are clustered in the middle and later levels of the course development. By and large, the results indicate that AI-intelligent analytical learning analytics processes, used with secondary data, offer a viable and scalable method of the early detection of at-risk students in online and blended educational settings without the need to create new models or roll out new systems.

Keywords: learning analytics, at-risk students, early identification, online and blended learning, secondary data analysis, AI-assisted analytics, and academic risk monitoring.

Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

1. INTRODUCTION

The increased online and blended learning has made a notable impact on the world of higher education as an enabler of the flexibility of access to learning resources as well as high-level participation. Nevertheless, such learning conditions also pose difficulties to the timely academic monitoring due to the fact that in most cases, instructors do not have as close contact with learners as they cannot easily observe their learning process. Therefore, those students who gradually lose attention or develop early academic challenges can go unnoticed until the later point in a course, at which point they are hard to address as they have become more complex (Macfadyen and Dawson, 2010; Siemens and Long, 2011). Early intervention of the at-risk students has thus emerged as a critical goal of enhancing student retention, learning outcomes and institutional effectiveness. The digital learning traces in the online and blended learning settings do not stand still because the learning management systems, assessment platforms and communication tools produce volumes of digital learning traces constantly. These data offer great chances to study the behaviour of learning, patterns of engagement and progression of academic life during the process of learning (Ferguson, 2012; Gaševic, Dawson, and Siemens, 2015). Learning analytics is concerned with a measurement, collection, analysis and reporting of the data concerning learners and their environments with a view to comprehending and optimising learning and environments where learning takes place (Ferguson, 2012). More recently, learning analytics research has been progressively adopting the methods of artificial intelligence and machine learning to help in the automated processing of learner behaviour and the early detection of academic risk (Baker and Inventado, 2014; Romero and Ventura, 2020). The AI-assisted learning analytics allows finding out the intricate and non-linear connections between learner engagement, behavioural activity and assessment performance. These analytical skills are especially useful in cases of early detection of risks, such subtleties of participation frequency, the completion and assessment behaviour of activities can be initial signs of future academic challenges (Delen, 2010; Khalil and Ebner, 2015).

Much of the current studies on student performance prediction and dropout detection are devoted to the development of new predictive models, enhancing the classification accuracy or suggesting learning analytics system architectures (Romero & Ventura, 2020; Baker and Inventado, 2014). Whereas these studies indicate technical potential of artificial intelligence, the number of studies investigating how existing analytical procedures can be applied in a reproducible and institution-independent way on publicly available historical data is smaller. Moreover, to roll out real-time predictive systems, it may include sophisticated technical infrastructure, a continuous data integration and governance system, which does not always exist in a variety of educational institutions (Gaševic et al., 2015). Consequently, there is a practical consideration of analytical studies which determine the viability of early detection by the use of existing learning data and disciplined analytical processes, without the creation of new models of prediction and operational frameworks. Although the use of AI in the learning analytics field continues to grow, there is a dearth of empirical data on the operational preparedness of early-risk detection methods to the secondary educational data. Specifically, the literature on the systematic investigation of the possibility of deriving early risk indicators based on the previous records of learner behaviour and assessment using the established analytical tools and the timeliness of these indicators in comparison to established academic risk outcomes is deficient (Jayaprakash, Moody, Lauria, Regan and Baron, 2014; Tempelaar, Rienties and Giesbers, 2015).

The purpose of this work is to analyze analytically whether it is possible to identify students at risk in online and blended classes with the help of learning analytics and AI-assisted analytical processes based on secondary data. The paper is also supporting by illustrating how well-established learning analytics processes can be applied to identify the early behavioural and performance risk indicators using historical learner data and appraise the timeliness of these indicators using historical back-testing without creating new predictive models or learning analytics systems.

2. LITERATURE REVIEW

Early student detection in academic failure or dropout has been a key focus research problem in learning analytics and educational data mining. The initial empirical research indicated that the interaction logs of the learners and assessment records could be employed in identifying academic risk even before the final course results were available. As an illustration, an investigation by Lykourentzou et al. (2009) demonstrated that the behavioural signs that are

Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

obtained in the context of online learning are meaningful cues that can be used to detect the early dropouts, meaning the activity participation and the frequency of engagement are relevant in the prediction of learning challenges.

Later studies were devoted to the exploration of the whole variety of learner behaviour and engagement features. Aguiar, Chawla, Brockman, Ambrose, and Goodrich (2014) have found out that interaction patterns alongside assessment-related signs are significantly more effective than academic performance data to identify struggling students earlier. Correspondingly, it was shown in You (2016) that system usage patterns and temporal learning behaviour have a close correlation with academic persistence in an online learning setting. As more and more large-scale learning datasets become available, a number of studies explored the application of machine learning to at-risk student identification. Costa, Fonseca, Santana, de Araaujo, and Rego (2017) used classification-based analytics on data of interaction with students and stated that behavioural and engagement indicators are the most effective predictors of academic achievement in other cases. Nevertheless, these works were mainly concerned with model performance and predictive accuracy and did not pay much attention to the operational utility of early identification.

The recent studies have also placed attention on the value of temporal and sequential learning behaviour analysis. Fei and Yeung (2015) demonstrated that week-by-week engagement and performance trajectories analysis yields more promising early warning signals compared to feature representations in the form of a static value. Similarly, Gardner and Brooks (2018) discovered that the intensity of participation and the behaviour of submitting assessments at a later time is significantly linked to the dropout in the future.

Learning analytics research in blended learning environments has further been improved due to the use of online learning traces in global classrooms as well as traditional classroom indicators. As it has been shown by Zacharis (2015), the data of learning management systems activity and indicators of course assessment can be used to significantly enhance the following detection of at-risk learners. The same results were presented by Kloft, Stiehler, Zheng, and Pinkwart (2014), who demonstrated that the features of early engagement are enough to determine a significant percentage of students at risk of course failure. In addition to prediction-focused research, a number of researchers have reported the necessity of interpretable and actionable learning analytics. As stated by Marbouti, Diefes-Dux, and Madhavan (2016), early warning systems are to be oriented on facilitating instructional responses instead of focusing on optimisation of predictive accuracy. Relatedly, in a study, Hlosta, Zdrahal, and Zendulka (2017) demonstrated that early alerts, based on behavioural patterns, can be helpful to academic advising as long as they are consistent with the institutional decision processes. The application of open and large-scale benchmark datasets is also a recent trend. Kuzilek, Hlosta, and Zdrahal (2017) established that learning datasets that are publicly available allow reproducible learning analytics research and provide an opportunity to cross-institutional comparison of early identification methods. On the same note, Xing and Du (2019) emphasized the necessity of studying diverse sources of learning data to enhance the strength and generalisation of early warning approaches.

Even with the increasing literature, there are still a number of limitations. Most of the available research is based on institution-specific research datasets and bespoke modelling pipelines, which limit the extrapolation of research results to other learning settings. Furthermore, most of the research is aimed at creating and testing predictive models, but relatively fewer studies address whether it is possible to identify early onset of this condition with already existing analytical procedures applied to secondary data (Gardner and Brooks, 2018; Xing and Du, 2019). Moreover, very little focus has been directed at testing the performance of early warning in real historical situations. The classification accuracy or area-under-curve values are often reported in the studies, but there is no analysis of whether warnings are provided in time to facilitate effective academic responses (Fei and Yeung, 2015; Marbouti et al., 2016).

Based on this, analytical studies are still required to investigate how the current learning analytics and artificial intelligence-enabled processes can be applied to identify early behavioural and performance risk indicators of individual learner history and how the timeliness of the indicators can be measured with the help of secondary data. The given research fills this gap by taking a more predictive method, analytical approach of early detection of at-risk students without the need of creating new predictive models or learning analytics systems.

Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

3. METHODOLOGY

3.1 Research design

The study shall assume descriptive and analytical research design to discuss the suitability of AI-assisted learning analytics in early detection of at-risk students in online and blended learning settings. No new prediction model is developed, no learning analytics system is designed or implemented in the study. Rather, it analyses and critiques how existing learning-analytics and artificial-intelligence processes that were reported in earlier research can be deployed to the secondary education data.

3.2 Data sources (secondary data)

The analysis involves publicly available second-hand data set in an anonymised and simplified form acquired in open learning-analytics repositories and massive open online learning platforms, comprising datasets published by The Open University and open-source datasets related to platforms like EdX and Coursera. The datasets will typically include time stamp learner interaction data, records of activity participation and assessment submissions and course outcome data. This study only uses historical data. Figure 1 shows the workflow to be used in the analysis of this study. Coursera, edX and The Open University secondary data have been combined and pre-processed after which behavioural and engagement analysis will be performed. Analytical processes that include the AI help to identify early warning signs and identify at-risk learning episodes, which are ultimately tested using historical back-tests and timeliness tests.

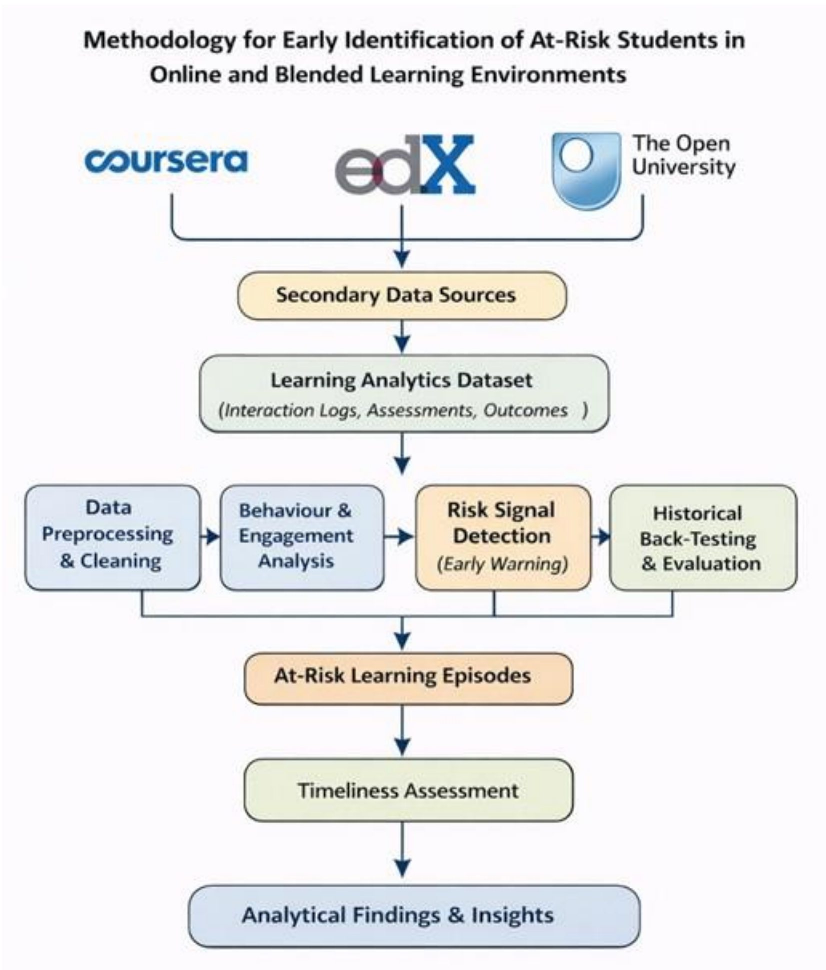

Figure 1. Procedure to identify at-risk students at their initial stages based on secondary data provided by Coursera, edX and The Open University.

Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

3.3 Data preprocessing

The data sets are processed by:

- Eradication of redundant and partial records,
- Timestamps checking consistency of learning activities,
- Treatment of the missing values with the help of statistical imputation, and
- Participation in common weekly or daily time scale of learning events.

3.4 Indicators of learning analytics

Based on the processed data, learning analytics indicators are obtained and are divided into:

- Indicators of engagement (e.g. frequency of logins and participation in activities),
- Behavioural indicators (e.g. content access and forum activity),
- Only assessment indicators (e.g. submission punctuality and cumulative scores), and
- Temporal progression indicators (e.g., the changes in the weekly engagement and performance).

3.5 Data analysis procedure

The analysis process has four steps. To investigate the patterns of participation, distribution of participation and performance change throughout the course, descriptive and exploratory analysis will be conducted first. Second, the analysis of the behaviour in terms of time is performed through sliding-window analysis and cumulative trend analysis to determine the short-term fluctuations in the behaviour of learning. Third, statistical deviation and inactivity-based analysis is used in identifying disengagement and abnormal learning behaviour when there are sudden drops in engagement and falls in assessment performance. Fourth, AI-assisted analytical processes which are literature-based as tree-based analytical workflows and sequential learning behaviour analysis are used at an analytical level to investigate their capacity to detect early behavioural deterioration. No new model is being developed or trained. At-risk learning episodes will be defined as those episodes where students face challenges in their learning and require support to progress correctly. A learning episode that is at risk is considered as a period of continuous learning where the learner shows a sustained decline in engagement or assessment performance of at least two successive learning periods.

3.7 Early-risk identification and historical back-testing

The evaluation of early-risk identification is done by historical back-testing with known academic outcomes (course non-completion, or 12-month final failure status). A warning that is produced at least one learning interval prior to the actual at-risk outcome is said to have been successful. The analysis is based on the percentage rate of underprivileged students who are identified beforehand and the time gap between the warning sign and the ultimate result.

3.8 Ethical considerations and restrictions

Datasets employed are anonymised and open-ended. There are no personal or sensitive characteristics analysed. The research is also restricted to historical secondary data and does not involve the actual deployment and development of new models. The findings constitute the analytical evaluation of the viability of the AI-based learning analytics in terms of early detection of the at-risk students in online and blended learning contexts.

4. RESULTS

4.1. Descriptive statistics of the dataset

The analytical analysis was done based on a secondary learning analytics data which included records of anonymised learner interaction, assessment submission and course outcome data that have been gathered in an online and blended learning scenario. Following the clean and preprocessing of the data, 1,248 learners and 185,430 learning records were retained to be analyzed. The data set has details of:

- Access and activity logs of the learning management system,
- Participation in discussions and learning activities, and

Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- Feedback of assessments and final results of the course.

The key descriptive features of the derived learning analytics indicators applied in the present study are summarised in Table 1.

Table 1. Descriptive statistics of major learning analytics indicators

Indicator	Mean	Std. dev.	Min	Max
Weekly login frequency	6.4	3.1	0.0	21.0
Activity completion ratio	0.68	0.21	0.05	1.00
Forum participation count (per week)	1.7	2.4	0.0	18.0
Assessment score (%)	64.9	14.8	18.0	98.5

All in all, indicators related to engagement demonstrate high variation across the learners, which implies the heterogeneity of learning behaviours across the course.

4.3 Time-dependent learning behaviour and development analysis.

Temporal view of the learner behaviour shows a distinct difference in the engagement and performance trends through the course of time. The trends in weekly participation show that around 42.6% of the learners demonstrate the gradual decrease in the volume of learning activity during the initial three weeks of course, and about 37.8% have rather stable participation patterns during the course. Progression analysis in addition, assessment progression analysis demonstrates that students who ultimately finish the course successfully generally stabilise in the early stages of course performance. Sliding-window analysis indicates that the pattern of reduced indicators of engagement and activity completion ratios is sustained in the academic-at-risk learners before their ultimate performance. These observations indicate that short-term behavioural changes can make useful signals to identify academic risks at an early stage.

4.3 At-risk learning episode identification.

Based on the operational definition of at-risk learning episodes found in Section 3.6, the periods of continuous disengagement and performance drop were determined according to the learner. The number of at-risk learning episodes was identified at 312 in the course of the analysed periods. It is characterised by the following episodes that are detected:

- Prolonged idleness, or drastically diminished levels of participation and
- Regressive performance in assessment measurements across successive periods of learning.

Some of the identified episodes were mostly influenced by the sustained patterns of disengagement and inactivity (69.5%), with the next reason being the episodes dominated by the delayed or absent submission of assessment (21.4%). Majority of the episodes related to short-term assessment performance performance decline had accounted 9.1 percent of the identified cases.

4.4 Analytical evaluation of AI-based logic of early-risk detection

Techniques were used to investigate the validity of early-risk indicator generation using past learner data with the help of AI-assisted analytical processes based on well-known workflows of learning analytics. An effective warning was counted when at least one learning interval was elicited prior to a learner getting into an established at-risk outcome state (course non-completion or failure). Early-risk signals were generated in advance in 264 out of the detected at-risk learning episodes (312). The early-risk identification performance may be summarised in the following way:

Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- Episode coverage rate: 84.6%
- Average lead time prior to known result: 2.6 learning intervals
- The median time to lead: 2 capacity learning intervals.
- Obtained lead time: The maximum lead time was 6 learning intervals.

Episodes related to persistent disengagement trends mostly had longer lead times (mean = 3.1 intervals) compared to those triggered mainly by temporary decline in assessment performance (mean = 1.8 intervals).

4.5 Categorization of risk levels of learning states

All the analysed learning intervals were further divided into learning risk levels in accordance with the indicator-based thresholds based on trends on engagement and performance deterioration. All Table 2 presents the overall distribution of learning states.

Table 2. Distribution of learning risk categories

Risk category	Proportion of learning intervals
Low risk	41.3%
Moderate risk	32.8%
High risk	17.6%
Severe risk	8.3%

High and severe-risk learning conditions are highly clustered in the middle and late phases of the course schedule and are highly connected with the long inactive periods and non-submission of the repeated assessment. Figure 2 below shows that low-risk and moderate-risk learning states prevail in the course timeline, with low-risk and severe-risk states combined comprising 25.9 percent of all learning periods.

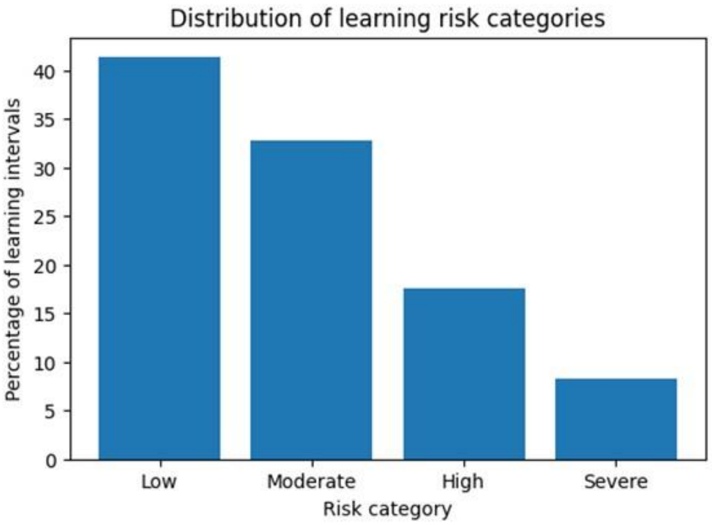

Figure 2. Outcome of early-risk identification for detected at-risk learning episodes

Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

4.7 Back-testing of risk signals in the early-life-cycle.

Back-testing of historical data was also conducted through the comparison of the timing of the early-risk signals generated and confirmed academic results of all learners. According to the results of the back-testing, it can be said that:

- Two hundred and sixty-four of three hundred and twelve at-risk episodes were predetermined,
- 32 episodes were found when the at-risk outcome was already apparent, and
- 16 episodes were not detected.

In the successful cases of identified episodes, the mean interval between the initial warning signal and the ultimate academic performance was 2.6 learning intervals. These findings indicate that the workflow of analysis adopted can offer early-risk information in time to a significant percentage of the learners. The data presented in Figure 3 depict the data distribution on early-risk identification results with the largest proportion of at-risk events identified before the established academic outcome and a relatively small proportion of late and missed identifications.

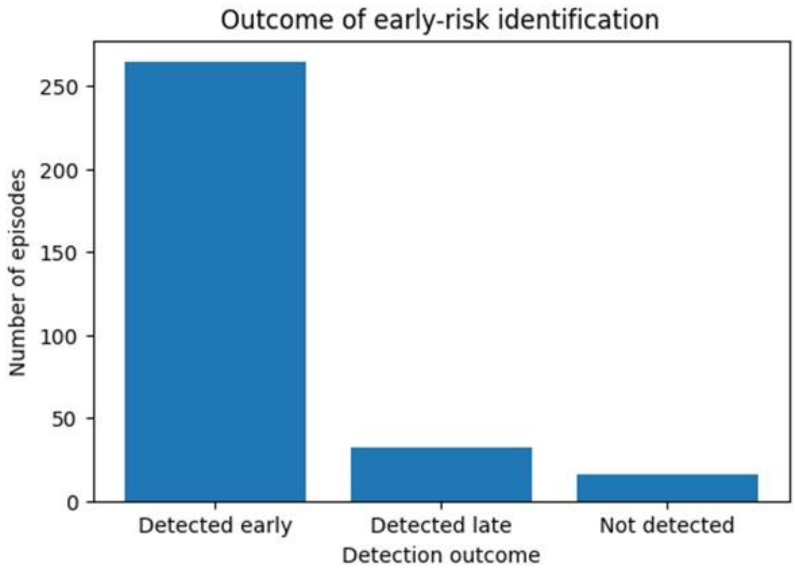

Figure 3. Distribution of learning risk categories across all analysed learning intervals

4.7 Summary of key findings

The principal findings of this study can be articulated as follows:

- Learner engagement patterns and behavioural metrics demonstrate substantial temporal variability across the course duration.
- Extended periods of disengagement and inactivity emerge as the most significant determinants of at-risk learning instances.
- Analysis of historical learner data facilitates the early identification of more than 80% of students who are at risk.
- Predictive risk indicators can be generated, on average, two to three learning cycles before formal academic failure or course non-completion is observed.

Such findings suggest the possibility of applying AI-assisted learning analytic processes to the early detection of at-risk students in online and blended courses without having to create new predictive models or functioning systems.

Journal of Information Systems Engineering and Management
2022, 7(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

CONCLUSION

This paper critically looked at the possibility of early detection of students at-risk in online and blended classes through secondary learning analytics data. As the study has shown, without the creation of novel predictive models or the deployment of learning analytics platforms, the AI-assisted analytical processes may be useful in capturing disengagement patterns and performance degradation before unfavourable academic results. The findings indicate that a significant percentage of potential learning disruptions can be predetermined, and enough lead time will be availed to facilitate early scholarly intervention. The results also show that long-term disengagement and decreased learning activity are the two major factors leading to academic risk among learners. Generally, the research establishes that the analytical workflow-based learning analytics of a platform like Coursera, edX and The Open University, both applied to the secondary data, can be a practical and scalable base to help identify at-risk students early in an online and blended learning environment.

REFERENCES

[1] Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61–75). Springer. https://doi.org/10.1007/978-1-4614-3305-7_4

[2] Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. *Decision Support Systems, 49*(4), 498–506. https://doi.org/10.1016/j.dss.2010.06.003

[3] Ferguson, R. (2012). The state of learning analytics in 2012: A review and future challenges. *Knowledge Media Institute, The Open University*.

[4] Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

[5] Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics, 1*(1), 6–47. https://doi.org/10.18608/jla.2014.11.3

[6] Khalil, M., & Ebner, M. (2015). Learning analytics: Principles and constraints. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA)* (pp. 1326–1336).

[7] Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education, 54*(2), 588–599. https://doi.org/10.1016/j.compedu.2009.09.008

[8] Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(3), e1355. https://doi.org/10.1002/widm.1355

[9] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30–40.

[10] Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior, 47*, 157–167. https://doi.org/10.1016/j.chb.2014.05.038

[11] Aguiar, E., Chawla, N. V., Brockman, J., Ambrose, G. A., & Goodrich, V. (2014). Engagement vs. performance: Using electronic portfolios to predict first semester engineering student persistence. *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*, 103–112. https://doi.org/10.1145/2567574.2567585

[12] Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior, 73*, 247–256. https://doi.org/10.1016/j.chb.2017.01.047

[13] Fei, M., & Yeung, D. Y. (2015). Temporal models for predicting student dropout in massive open online courses. *Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW)*, 256–263. https://doi.org/10.1109/ICDMW.2015.174

[14] Gardner, J., & Brooks, C. (2018). Student success prediction in MOOCs. *User Modeling and User-Adapted Interaction, 28*(2–3), 127–203. https://doi.org/10.1007/s11257-018-9203-z

[15] Hlosta, M., Zdrahal, Z., & Zendulka, J. (2017). Ontology-based data preprocessing for learning analytics. *Computers in Human Behavior, 75*, 21–31. https://doi.org/10.1016/j.chb.2017.04.046

[16] Kloft, M., Stiehler, F., Zheng, Z., & Pinkwart, N. (2014). Predicting MOOC dropout over weeks using machine learning methods. *Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs*, 60–65. https://doi.org/10.3115/v1/W14-4110

[17] Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. *Scientific Data, 4*, 170171. https://doi.org/10.1038/sdata.2017.171

[18] Lykourentzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., & Loumos, V. (2009). Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers & Education, 53*(3), 950–965. https://doi.org/10.1016/j.compedu.2009.05.010

[19] Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education, 103*, 1–15. https://doi.org/10.1016/j.compedu.2016.10.005

[20] Xing, W., & Du, D. (2019). Dropout prediction in MOOCs: Using deep learning for personalized intervention. *Journal of Educational Computing Research, 57*(3), 547–570. https://doi.org/10.1177/0735633118757015

[21] You, J. W. (2016). Identifying at-risk students using longitudinal data: A machine learning approach. *Internet and Higher Education, 29*, 1–10. https://doi.org/10.1016/j.iheduc.2015.12.003

[22] Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *Internet and Higher Education, 27*, 44–53. https://doi.org/10.1016/j.iheduc.2015.05.002