*Environment. Technology. Resources. Rezekne, Latvia*
*Proceedings of the 16th International Scientific and Practical Conference. 2025. Volume II, 57-64*

# *Methodology for Analysing LMS Data to Predict Student Dropout Risk in Higher Education*

**Linda Luīze Barbare**
*Institute of Applied Computer Systems*
*Riga Technical University*
Riga, Latvia
linda-luize.barbare@rtu.lv

**Aleksejs Jurenoks**
*Institute of Applied Computer Systems*
*Riga Technical University*
Riga, Latvia
aleksejs.jurenoks@rtu.lv

**Magone Una Rauba**
*Institute of Applied Computer Systems*
*Riga Technical University*
Riga, Latvia
magone-una.rauba@rtu.lv

**Zane Višķere**
*Institute of Applied Computer Systems*
*Riga Technical University*
Riga, Latvia
zane.viskere@rtu.lv

***Abstract*—Nowadays, educational institutions use Learning Management Systems (LMS) to support students in the learning process. LMS technical data analysis enables the monitoring of student activities and early identification of those at risk of failing a course. Data gathered during the educational process facilitates the adaptation of learning content to meet each student's individual needs. By leveraging this data, institutions can implement adaptive education, allowing study programs to be structured based on personalized learning pathways, intelligent recommendation systems, and dynamic curriculum adjustments. Additionally, by analyzing student model data, it is possible to assess dropout risks. As a result, research on student attrition rates has gained increased attention. This paper examines the methodology for analyzing Moodle LMS data to adaptively detect factors influencing student dropout risk. The research explores the potential of analyzing log file data generated by Moodle LMS to identify student model parameters and their impact on student success throughout the entire educational process. By utilizing learning patterns and engagement indicators, activity log data from more than seven hundred students at Riga Technical University's Moodle e-learning system was analyzed. The research aimed to identify correlations and relationships between several factors, including the availability of resources for students, the number of graded activities, activity types, views, and other relevant data. By analyzing correlations between fluctuations in students' learning achievements and behavioral patterns in e-learning platforms, the study aims to identify key indicators and metrics for predicting dropout tendencies. The findings suggest that a decline in engagement, the presence of negative patterns, or the absence of consistent learning behaviors serve as reliable indicators of students at risk of dropping out.**

***Keywords — learning management system, student dropout, e-learning, education theories***

## I. INTRODUCTION:

The growing availability of educational data has an increasing topicality in Educational Data Mining (EDM) and Learning Analytics (LA) [1] – [4]. LMSs platforms such as Moodle generate large volumes of data related to student activities. Analysing this data can reveal patterns in student engagement and predict dropout risks. [4], [5]. The goal of this paper is to describe a methodology of analysing Moodle LMS logs to identify factors influencing student dropout risk and to support early intervention strategies. [1], [4], [5]. Analysing student activity logs in LMS is a non-invasive method that relies on behaviour-related data to predict student success by examining student interactions within the LMS, such as access times, resource utilization, and participation in activities, without requiring personal or sensitive data [1].

RQ1: Does LMS log data correlate with pedagogical and cognitive aspects on e-learning?

RQ2: What specific behavioural patterns in Moodle LMS log data are most indicative of a heightened dropout risk?

Studies prove that low engagement metrics such as time spent on tasks, participation in discussions and

*Online ISSN 2256-070X*
*https://doi.org/10.17770/etr2025vol2.8613*

forums, and interaction with course materials in online courses is a strong indicator of student dropout [1], [2]. Even more, it has been concluded, that students who perform poorly in the early stages of a course are at higher risk of discontinuing their studies [6], [7] Majority of studies focus on log activity metrics but do not focus on identifying reasons for student disengagement, leaving a gap in understanding the factors that impact participation and motivation in online learning environments.

*Related studies:*

A systematic literature review was conducted to identify factors related to student dropout. To ensure the exploration of the latest methodologies in this research domain, particularly considering the increased adoption of online learning due to the COVID-19 pandemic, studies published between 2019 and 2025 were selected. The literature search was conducted by using keywords: ***learning management system (LMS), student dropout, e-learning, education theories.***

Initial search string "dropout AND Moodle OR learning AND management AND system AND prediction" returned 141 papers in Scopus that were narrowed down to 75 after adjusting selection criteria (in English, peer-reviewed articles, relevance to LMS and dropout prediction, etc.) After reviewing titles and abstract, 30 articles were identified to be directly relevant to research domain. Studies that have been included in review were focusing on Higher education institutions (HEI) and can be categorized by 2 main topics explored: machine learning techniques and models in dropout prediction and engagement metrics and key factors influencing dropout.

Student dropout has been critical research area for HEI as it directly impacts institutions' financial health and is a key aspect to foster society's progress and economic development [6], [7], [12], [13]. Recently there has been an increase of number of studies where student dropout risk is predicted using computer science and machine learning technologies [1], [7] – [11], [13], [14]. It has been identified, that in research domain related studies machine learning techniques such as Decision Trees, Random Forests, Logic Regressions, K-Nearest Neighbour (KNN), Clustering, Neural Networks, etc., are widely used to analyse LMS data for real-time data monitoring to track students' activity and interaction with course content providing valuable insights on student performance [1] – [17]. As established in literature, multiple tools have been proposed for dropout prediction, however many of tools in market tend to be too general and are not modified accordingly to educational institution specifics – such as teaching field, course length, internal organization policies and grading [7], [15] Increased adaptation of online and mixed-learning environments has highlighted the need for more advanced, AI-driven tools to support student success [7], [15]

Many studies use combined machine learning technique models and deep learning algorithms to achieve higher accuracy rates. The ensemble model by Rabelo and Zarate combines results from Logistic Regression, Neural Networks, and Decision Tree techniques developing a conceptual model that is primary based on previous knowledge – surveys from dropout students, knowledge from literature review from to domain related works. It was able to predict dropouts with 98.1% accuracy [7]. Similar study done by Tamada, Gusti, et al, compares seven machine learning algorithms (Random Forest, Decision Tree, Gradient Boosted Trees, Logistic Regression, Naive Bayes, Deep Learning, and Support Vector Machine) on analysing 761 high school student data at 20%-40%-60% course completion stages total across 8 classes within a 2-year period. Highest accuracies were achieved using methods Deep Learning (90.33% at 60% course completion stage) and Random Forest (88.42% at 40% course completion stage) [16].

Framework introduced by Shoaib and Sayed proposes combining pre-processing techniques using Convolutional Neural Network (CNN) for hidden pattern extraction with ensemble model to implement AI-driven education technologies. The model was able to categorize students into "No Risk," "Low," "Moderate," and "High" risk levels and predict the risk of dropout with 93% accuracy. Results were validated against actual student outcomes, demonstrating its reliability in real-world scenarios [6]. Another study analysed 12 000 log file entries using CNN [12]. Study suggested that machine learning classifiers alone do not provide efficient mechanisms on student performance prediction and achieving high accuracy with a large data set is challenging, however proposed model scored average accuracy of 98.6% [12]. The study by Sáiz-Manzanares and Rodríguez-Díez was conducted on single course with 49 students as a result the UBUMonitor tool was developed – a desktop app that connects to Moodle via REST API and WEB services to visualise log activities [15].

Studies propose various key metrics that impact dropout rates, including behavioural and engagement patterns, socio-demographic factors, teaching methods, and student personality traits [10]. These insights provide an understanding of how different factors impact dropout rates across various HEIs and regions. Study by Al-Kindi, Al-Khanjari, et al, was conducted in Sultan Quabos university and analysed records from "Introduction to computer science course" with 29 students (in total 273 906 log files) within semester. Four main key factors were extracted: Engagement (E), behaviour (B), Personality (Pers.) and Performance (P) to identify specific patterns. EBPersP model concluded that the most influential factor on performance was personality, followed by behaviour and engagement, implicating that understanding the student's personality traits can lead to more effective teaching strategies. [5] In research [13], it was established that older students tend to drop out more, but in [12] a strong correlation for higher drop out risk was found for students between the age of 20-29. However, the HEI type differed – research [12] was conducted in Arab Online University, while [13] was executed in mixed type HEI

(Face to face (F2F) and massive open online courses (MOOC)) France.

Building a ML model for student performance prediction necessitates the collaborative efforts of specialists in education, digital curriculum design, applied statistics, and computer science. László Bognár and Tibor Fauszt involved experts in these fields and found that introducing chapter-level indicators can increase the accuracy of predictions compared to using only course-level or site-level indicators [17].

Most of the research studies focus on Moodle logs activity and engagement as behavioural metrics, but do not specifically underline the reasons for student disengagement. Research is mainly conducted on engagement activity in combination with demographic or financial factors but exclude feedback from students and educators.

*Pedagogical and cognitive aspects on e-learning*

Online learning offers the advantage of education accessibility and flexibility [1] – enabling students to engage with any available online course at any time from any location. This enhances learning opportunities, while also is beneficial to the educators by expanding target audience reach and optimizing resource utilization. However, stable internet connection and access to digital devices has been identified as restrictive challenge in terms of e-learning accessibility [19]. Digital learning is a sustainable approach to global education challenges providing equity and inclusivity by ensuring digital connectivity and quality content delivery [18].

The increasing integration of technologies in education has prompted digital pedagogy methods to shift from traditional pedagogical frameworks such as Vygotsky, Piaget, etc., and calls for more student-centred approach [19] encouraging students to actively participate in knowledge construction [18], [20]. However, dependency on students' self-regulated learning that allows learners to navigate their progress without direct supervision and peer support [19] might result in cognitive overload and reduced engagement [18], [21], [22] that leads to higher dropout rates [8] and low academic scores [19] in comparison to traditional face to face classes. Therefore, it is crucial to timely allocate risk factors to provide additional support and mitigate dropout rates [20].

The central key to academic success and building learning experiences is effective student interaction with education content in an e-learning environment. [21] Pedagogically, e-learning environments encourage active participation through various interactive methods such as collaborative technologies, gamification, forums, etc., and promote self-regulated learning allowing students to guide their education process based on learning style and preferences through digital platforms [19], [21]. Cognitively, e-learning environments must provide deep learning experiences that can be achieved through content personalization and meaningful interactions rather than surface-level engagement [18], [21]. In the year 1989 Moore described three types of interaction: Learner-Teacher (LT), Learner-Content (LC), and Learner-Learner (LL) which outlines forms of engagement in an online environment. The Community of Inquiry (CoI) framework, which was influenced by Moore's research, emphasizes the importance of cognitive presence, where students can actively engage in critical thinking and problem-solving rather than passive content consumption. Furthermore, the Self-Determination Theory (SDT) by Yuerong and Na (2024) emphasizes that personal motivation and engagement in e-learning strongly correlate with three psychological needs: autonomy, competence, and relatedness [19], [21]. The Adaptive Structural Learning and Technology Acceptance Model (ASL-TAM) by Zheng, Lou, et al, also emphasizes students' relatedness to the topic and identifies three key dimensions – perceived ease of use, perceived usefulness, and system adaptability – as critical factors that shape student engagement and satisfaction with online learning platforms [22]. Another element that has a crucial role in student engagement levels is interactive learning methods such as forums, problem-solving tasks, adaptive learning pathways, etc., that foster cognitive presence [20], [21]. Scaffolding strategies, such as guided inquiry and structured feedback loops promote an increase in cognitive engagement and can assist students in structuring knowledge through reflective discourse [21]. If e-learning environments fail to provide meaningful scaffolding, it can result in students developing misconceptions and disengagement from further learning process [21]. Effective e-learning environments must combine interactive strategies, multimedia resources, and frequent engagement through forums and assignments to improve performance and reduce dropout [21]. However, it is suggested, that if there is an absence of direct social and teaching presence from educators in online learning courses, that can be mitigated through artificial intelligence-driven analytics that provide real-time feedback and personalized learning recommendations [21].

The studies show that student engagement varies based on multiple factors not only including students' personal traits and preferences but also digital accessibility and online course content and instructional quality [19]. Flexibility in online learning environments can enhance academic performance, however, challenges such as reduced peer and educator interaction, digital and lack of motivation can negatively affect student engagement and therefore increase dropout rates [19]. It is suggested that online learning environment courses should implement interactive strategies, multimedia resources, and frequent LT, LC, and/or LL interactions through various digital collaboration tools to enhance both performance and engagement [19], [21].

While e-learning provides adaptability and accessibility, learning outcomes depend on pedagogical design and used methods, cognitive support strategies, and

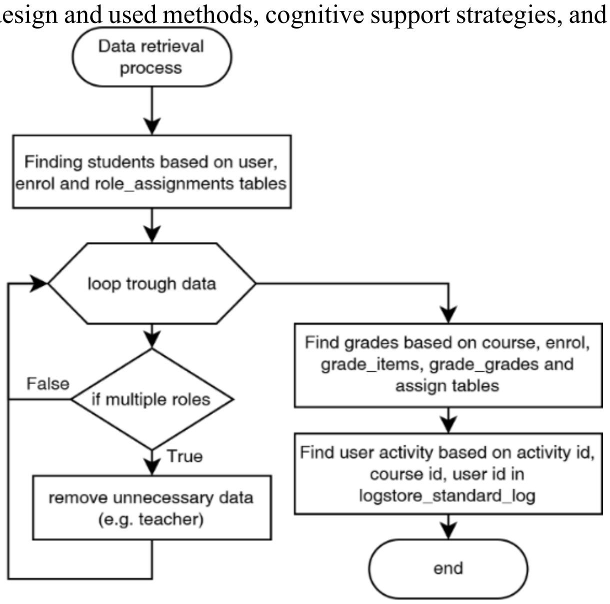


Fig. 1. Data retrieval process.

interaction strategies. Aligning e-learning environments with digital pedagogy strategies would result in enhanced learning outcomes and reduced dropout rates [18] – [22].

## II. MATERIALS AND METHODS

Data mining is the process of discovering new knowledge in large amount of already existing information, this process is based on traditional statistical analysis, mathematics and machine learning [23]. This process supports data driven decision making which in turn supports predictive analytics based in historic data by use of classification, regression or time series [23]. Predictive analytics uses past data to forecast future events, based on patterns extracted from data [23]. In this research, time series are used as base for predictions.

During data mining Moodle data, multiple steps had to be taken to achieve the results – data retrieval, data preparation, model building and pattern mining and analysing [23]. During the retrieval of data, data is mainly selected from logstore_standard_log (that contains course element views and corresponding user IDs), User (contains Users name, surname and contact information), Enrol (contains user enrolment in course date), Course (contains course name, course start and end date), role_assignments (contains users assigned roles for a course), grade_items (contains information about if assignment contains multiple grades), grade_grades (contain individual assignment grades, MIN and MAX assignable grade), Assign (contains assignment course information, Assignment name, deadlines) tables [24].

Due to large data amount in this particular case logstore_standard_log table (which is part of events API) is stored in separate data base together with other event related tables as per industry standard for cases when information needs to be split between multiple databases [25] – [27]. This made the data retrieval a multistep process (Fig. 1). First, all students who learn in a particular semester are found based on course, enrol, and user tables, and then their data is found in their respective databases to limit unnecessary data retrieval by using WHERE clauses. In further preparation process data is cleaned from data that could not be efficiently filtered out during the initial retrieval due to performance issues such as reaching environments timeout limit for single request length, cleaning included unnecessary log line removal (e.g., general course page views), duplicate records are removed or consolidated by regrouping to contain as little duplicate information as possible [28].

A two-week period ($t_{learn}$) was used to define participation windows. [29], [30]. Based on learning time frame length ($t_{learn\ len}$,(1)), student activities are looked at in time frames, which created one of the indicators for student non-participation ($t_{np\ activity}$). Students missing multiple time frame ($t_{learn}$), can be considered at high risk of dropping out, as knowledge retention (forgetting curve) needs to be considered. Additionally, the curriculum life span evaluation ($t_{buffer}$ and $t_{u\ buffer}$) must be performed, as missing multiple learning periods may make it impossible to adjust the learning plan to the student's needs, potentially leading to dropout ((2), (3), comparison (4)) [29], [30]. Next step is to look at students grades more closely, where a conclusion that, if student has used up all learning process buffer (5) they are at a higher risk for not completing this subject, is made, as well as if their overall grades are under passing grade increased by standard deviation they are at high risk of dropping the course as additional improvements to grades are needed ($t_{impr}$, (6), (7), (8), comparison (9)) [29].

TABLE 1 KEY VARIABLES

| Variable | Description |
|---|---|
| $t_{course\ len}$ | Length of course in weeks |
| $t_{activity}$ | Total activity length |
| $t_{learn\ len}$ | One learning time frame length |
| $t_{buffer}$ | Course buffer time, based on course length and a constant |
| $t_{np\ acitivity}$ | Non-participated activity count, in other words, times not participated |
| $t_{u\ buffer}$ | Used buffer time for non-participated activities |
| $t_{lo\ buffer}$ | Leftover buffer time based on time needed to finish activities |
| $\sum grade$ | Sum of all students' grades |
| $grade\ count$ | Student grade count |
| $avg\ grade$ | Average grade for student |
| $std\ dev$ | Standard deviation |
| $passing\ grade$ | Grade to pass the course |
| $t_{impr}$ | Time needed to improve grades |
| $t_{upd\ buffer}$ | Updated buffer value based on needed improvements |
| $\Delta val$ | Difference between students' average grade and average grade in course |
| $avg\ pos$ | Average position |
| $\Delta pos$ | Difference between students grade position and average position in course |
| $t_{corr}$ | Buffer time correction |
| $t_{f\ buffer}$ | Final buffer time |

During experiments additional corrections were made based on student grade rating ((10), (11), comparison (12)) in the particular course to increase the precision of prediction, so that there are less false positives for low-risk students (Fig. 2). These initial findings created a prediction of how likely the student is to finish a certain course, after which based on course information (e.g, ECTS) additional adjustments could be made to adjust for course importance.

The table (Table 1) defines key variables used in the dropout prediction model, covering time-based metrics for course engagement, buffer calculations and grade-based indicators for assessing student performance. These variables enable the model to evaluate learning participation patterns and identify students at risk of dropping out.

Additionally, to methodology we need to look at the calculations that make these predictions possible.

$$t_{learn\ len} = \frac{t_{course\ len}}{t_{activity}} \quad (1)$$

(1) Describes calculation of course learning time frame length. Where $t_{learn\ len}$ is one learning time frame length, $t_{course\ len}$ is the total course length, $t_{activity}$ is total course activity length.

$$t_{buffer} = t_{course\ len} \cdot 0.3 \quad (2)$$

(2) Represents course buffer time calculation, where 0.3 is course buffer time constant based on research [29], [30].

$$t_{u\ buffer} = \frac{t_{course\ len}}{t_{activity}} \cdot t_{np\ acitivity} \quad (3)$$

(3) Calculated used buffer time, where $t_{np\ acitivity}$ is non-participated activity count.

$$if\ t_{u\ buffer} > t_{buffer}\ then\ student\ is\ hight\ risk \quad (4)$$

(4) Shows the comparison used to determine high risk students based on buffer usage (1) (2) (3).

$$t_{lo\ buffer} = t_{buffer} - t_{u\ buffer} \quad (5)$$

(5) Calculates leftover buffer based on non-participated activity count. This shows current situation per student on already predicted need for learning process buffer usage.

$$avg\ grade = \frac{\sum grade}{grade\ count} \quad (6)$$

(6) Calculates average grade for this student.

$$t_{impr} = (passing\ grade + std\ dev - avg\ grade) \cdot t_{learn\ len} \quad (7)$$

(7) Calculates the buffer time necessary to improve grades, where passing grade is course passing grade.

$$t_{upd\ buffer} = t_{lo\ buffer} - t_{impr} \quad (8)$$

(8) Updates leftover buffer based on passing grade.

$$if\ t_{impr} > 0\ then\ if\ t_{upd\ buffer} > t_{buffer}\ then\ student\ is\ hight\ risk \quad (9)$$

(9) Shows comparison for high risk based on need for improvements.

$$t_{corr} = \Delta val \cdot \frac{\Delta pos}{avg\ pos} \cdot t_{learn\ len} \quad (10)$$

(10) Shows experimentally determined formula for correction value for buffer.

$$t_{f\ buffer} = t_{upd\ buffer} - t_{corr} \quad (11)$$

(11) Calculates buffer after correction.

$$\begin{aligned}
&if\ |pos_{max} - pos_{stud}| > |pos_{min} - pos_{stud}| \\
&then \\
&\quad if\ avg\ grade < passing\ grade + std\ dev \\
&\quad then \\
&\qquad if\ avg\ grade_{pos} < course\ avg\ grade_{pos} \\
&\qquad then \\
&\qquad\quad Formula\ 10\ and\ Formula\ 11\ and \\
&\qquad\quad if\ t_{f\ buffer} < 0 \\
&\qquad\quad then\ student\ is\ hight\ risk \\
&else\ no\ value\ correction
\end{aligned} \quad (12)$$

(12) Compares whether student grades indicate that resulting buffer should be corrected by using (10) and (11) (Fig. 2).

Validation process used Riga Technical University Faculty of Computer Science, Information Technology and Energy data from learning year 2023/2024 autumn semester. Experiments main task was to check the dropout risk based on data available in e-learning platform. Data was collected with 1 month intervals.

During the experiment, predictive model determined risk factor which was then compared to actual situation based on semester results. The results are categorised in three categories – predictive results match the outcomes, predictive results show a false positive (student should pass the course, but they don't), predictive results show a false negative (student should not pass, but they pass the course).

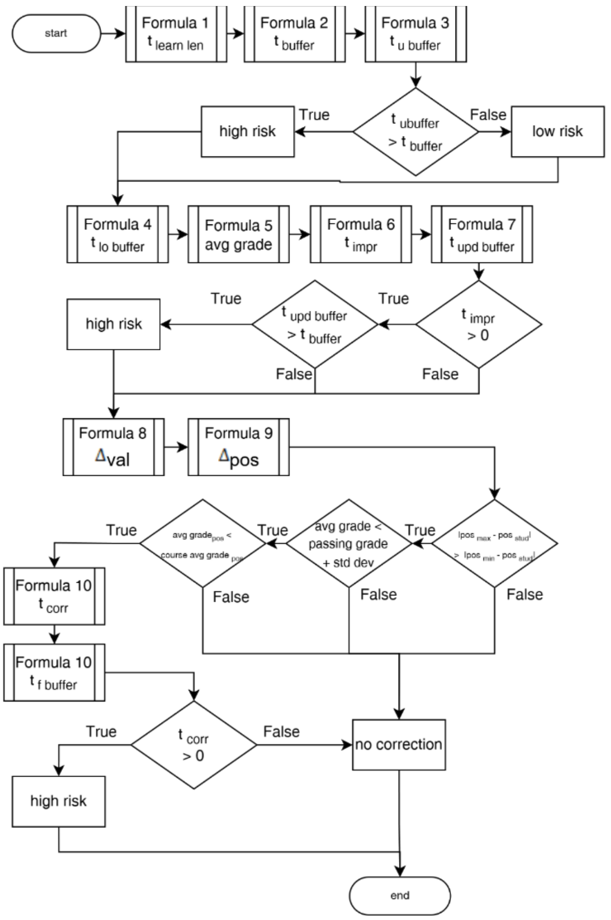


Fig. 2. Prediction process.

## III. RESULTS AND DISCUSSION

In first iteration, near the start of the course, the results 41% matched, 38% showed false positive and 21 % false negative. Second iteration determined 42% accurately, 40 % false positive and 18% false negative. Third iteration showed significant improvement in accuracy with 64% accurate predictions, 22% false positive and 14% false negative. In last iteration near the end of course the method showed 76% accurate results, 16% false positive and 8% false negative.

Based on during the experiment acquired data, it is possible to see that the method can identify course fail probability when 50% of the course has passed. Experiment showed that there are situations when it' is not possible to identify drop out risk without knowing study course complete requirements (e.g. false positive case context – it is not possible to identify the cause without knowing the requirements e.g. exam or activity weights). As a result, the research proves that just using LMS technical data the model can identify risk groups, but it's necessary to develop study course activity criteria (coefficients) that would help in model decision making.

The model's effectiveness is limited by its reliance on complete and consistent LMS data as well as set course curriculum and cannot account for unrecorded external influences such as personal or motivational factors. Enhancing model precision will require the development of course-specific coefficients and ECTS.

## IV. CONCLUSIONS

The increasing adoption of digital education technologies, particularly e-learning environments like Moodle, have transformed the way how the student activities and engagement are being analysed. By examining log file data, institutions can early identify at-risk students and provide personalized support such as individual learning plans, additional support, etc., to mitigate dropout levels.

This research analysed activity log data from over 700 students at Riga Technical University's Moodle e-learning environment and focused on identifying key engagement metrics and patterns that corelate with dropout tendencies, aiming to provide actionable insights for educators for timely intervention.

The study proposes a data-driven methodology for analysing Moodle log data that combines analysis of time-series with buffer time calculations and grade-based corrections to predict dropout risks.

Findings demonstrate that low engagement – non-participation within defined learning time frame ($t_{learn}$) – and low grades below to passing thresholds are strong predictors of potential dropout. These findings align with previous results from to domain related studies highlighting the importance of early engagement and performance analytics in online courses. This research concludes that decline in engagement and inconsistency in participation strongly corelates with negative academic performance and risk of dropping out of course. To mitigate potential risks, timely intervention based on engagement analytics is crucial.

Future research should focus on understanding the reasons behind decline in student engagement in context of course content quality and terms of pedagogical and cognitive aspects. Additionally, further investigation on the long-term impact of digital pedagogy strategies on learning outcomes is necessary. By implementing continuous improvements and aligning online course content with digital pedagogy strategies, HEIs can bridge the gap between technology and student success, fostering sustainable education and improve retention rates.

## V. ACKNOWLEDGMENTS

This study was partly supported by Project No. 5.2.1.1.i.0/2/24/I/CFLA/003 –Implementation of consolidation and governance changes at Riga Technical University, LiepU, RTA, LJA, and LJK for advancement towards excellence in higher education, science, and innovation.

## REFERENCES

[1] J. Pecuchova and M. Drlik, "Enhancing the Early Student Dropout Prediction Model Through Clustering Analysis of Students' Digital Traces," IEEE Access, vol. 12, pp. 159336–159367, 2024, [Online] Available: IEEE Xplore.

https://ieeexplore.ieee.org/document/10736593 [Accessed Mar. 04, 2025.], https://doi.org/10.1109/ACCESS.2024.3486762

[2] I. D. Stanciu, Á. Hernández-García, M. Á. Conde, and N. Nistor, "Decoding a decade. Trends and evolution in learning analytics: A comprehensive synthesis," Computers in Human Behavior, vol. 165, p. 108526, Apr. 2025, [Online] Available: ScienceDirect https://www.sciencedirect.com/science/article/pii/S0747563224003947?via%3Dihub [Accessed Mar. 04, 2025.], https://doi.org/10.1016/j.chb.2024.108526

[3] S. Sarker, M. K. Paul, S. T. H. Thasin, and Md. A. M. Hasan, "Analyzing students' academic performance using educational data mining," Computers and Education: Artificial Intelligence, vol. 7, p. 100263, Dec. 2024, [Online] Available: ScienceDirect https://www.sciencedirect.com/science/article/pii/S2666920X24000663?via%3Dihub [Accessed Mar. 04, 2025.], https://doi.org/10.1016/j.caeai.2024.100263

[4] I. Mwalumbwe and J. Mtebe, "Using Learning Analytics to Predict Students' Performance in Moodle Learning Management System: A Case of Mbeya University of Science and Technology," The Electronic Journal of Information Systems in Developing Countries, vol. 79, no. 1, pp. 1–13, Mar. 2017, [Online] Available: Wiley Online Library https://onlinelibrary.wiley.com/doi/10.1002/j.1681-4835.2017.tb00577.x [Accessed Mar. 04, 2025.], https://doi.org/10.1002/j.1681-4835.2017.tb00577.x

[5] I. Al-Kindi, Z. Al-Khanjari, and Y. Jamoussi, "Extracting student patterns from log file Moodle course: A case study," International Journal of Evaluation and Research in Education (IJERE), vol. 11, no. 2, Art. no. 2, Jun. 2022, [Online] Available: IJERE https://ijere.iaescore.com/index.php/IJERE/article/view/23242 [Accessed Mar. 04, 2025.], https://doi.org/10.11591/ijere.v11i2.23242

[6] M. Shoaib, N. Sayed, J. Singh, J. Shafi, S. Khan, and F. Ali, "AI student success predictor: Enhancing personalized learning in campus management systems," Computers in Human Behavior, vol. 158, p. 108301, Sep. 2024, [Online] Available: ScienceDirect https://www.sciencedirect.com/science/article/abs/pii/S0747563224001699?via%3Dihub [Accessed Mar. 04, 2025.], https://doi.org/10.1016/j.chb.2024.108301

[7] A. M. Rabelo and L. E. Zárate, "A model for predicting dropout of higher education students," Data Science and Management, vol. 8, no. 1, pp. 72–85, Mar. 2025, [Online] Available: ScienceDirect https://www.sciencedirect.com/science/article/pii/S2666764924000341?via%3Dihub [Accessed Mar. 04, 2025.], https://doi.org/10.1016/j.dsm.2024.07.001

[8] H. S. Park and S. J. Yoo, "Early Dropout Prediction in Online Learning of University using Machine Learning," JOIV : International Journal on Informatics Visualization, vol. 5, no. 4, pp. 347–353, Dec. 2021, [Online] Available: JOIV https://joiv.org/index.php/joiv/article/view/732 [Accessed Mar. 04, 2025.], https://doi.org/10.30630/joiv.5.4.732

[9] M. M. Tamada, R. Giusti, and J. F. de M. Netto, "Predicting Students at Risk of Dropout in Technical Course Using LMS Logs," Electronics, vol. 11, no. 3, Art. no. 3, Jan. 2022, [Online] Available: MDPI https://www.mdpi.com/2079-9292/11/3/468 [Accessed Mar. 04, 2025.], https://doi.org/10.3390/electronics11030468

[10] Z. Song, S.-H. Sung, D.-M. Park, and B.-K. Park, "All-Year Dropout Prediction Modeling and Analysis for University Students," Applied Sciences, vol. 13, no. 2, Art. no. 2, Jan. 2023, [Online] Available: MDPI https://www.mdpi.com/2076-3417/13/2/1143 [Accessed Mar. 04, 2025.], https://doi.org/10.3390/app13021143

[11] M. Vaarma and H. Li, "Predicting student dropouts with machine learning: An empirical study in Finnish higher education," Technology in Society, vol. 76, p. 102474, Mar. 2024, [Online] Available: ScienceDirect https://www.sciencedirect.com/science/article/pii/S0160791X24000228?via%3Dihub [Accessed Mar. 04, 2025.], https://doi.org/10.1016/j.techsoc.2024.102474

[12] ]M. Sayed, "Student Progression and Dropout Rates Using Convolutional Neural Network: A Case Study of the Arab Open University," Journal of Advanced Computational Intelligence and Intelligent Informatics, vol. 28, no. 3, pp. 668–678, May 2024, [Online] Available: ResearchGate https://www.researchgate.net/publication/380726606_Student_Progression_and_Dropout_Rates_Using_Convolutional_Neural_Network_A_Case_Study_of_the_Arab_Open_University [Accessed Mar. 04, 2025.], https://doi.org/10.20965/jaciii.2024.p0668

[13] M. Phan, A. De Caigny, and K. Coussement, "A decision support framework to incorporate textual data for early student dropout prediction in higher education," Decision Support Systems, vol. 168, p. 113940, May 2023, [Online] Available: ScienceDirect https://www.sciencedirect.com/science/article/abs/pii/S0167923623000155?via%3Dihub [Accessed Mar. 04, 2025.], https://doi.org/10.1016/j.dss.2023.113940

[14] ]V. Realinho, J. Machado, L. Baptista, and M. V. Martins, "Predicting Student Dropout and Academic Success," Data, vol. 7, no. 11, Art. no. 11, Nov. 2022, [Online] Available: MDPI https://www.mdpi.com/2306-5729/7/11/146 [Accessed Mar. 04, 2025.], https://doi.org/10.3390/data7110146

[15] M. C. Sáiz-Manzanares, J. J. Rodríguez-Díez, J. F. Díez-Pastor, S. Rodríguez-Arribas, R. Marticorena-Sánchez, and Y. P. Ji, "Monitoring of Student Learning in Learning Management Systems: An Application of Educational Data Mining Techniques," Applied Sciences, vol. 11, no. 6, Art. no. 6, Jan. 2021, [Online] Available: MDPI https://www.mdpi.com/2076-3417/11/6/2677 [Accessed Mar. 04, 2025.], https://doi.org/10.3390/app11062677

[16] ]M. Tamada, R. Giusti, and J. Magalhaes Netto, Predicting Student Performance Based on Logs in Moodle LMS. p. 8. 2021, [Online] Available: IEEE Xplore https://ieeexplore.ieee.org/document/9637274 [Accessed Mar. 04, 2025.], https://doi.org/10.1109/FIE49875.2021.9637274

[17] L. Bognár and T. Fauszt, "Factors and conditions that affect the goodness of machine learning models for predicting the success of learning," Computers and Education: Artificial Intelligence, vol. 3, p. 100100, Jan. 2022, [Online] Available: ScienceDirect https://www.sciencedirect.com/science/article/pii/S2666920X22000558?via%3Dihub [Accessed Mar. 04, 2025.], https://doi.org/10.1016/j.caeai.2022.100100

[18] R. Huang, M. A. Adarkwah, M. Liu, Y. Hu, R. Zhuang, and T. Chang, "Digital Pedagogy for Sustainable Education Transformation: Enhancing Learner-Centred Learning in the Digital Era," Front. Digit. Educ., vol. 1, no. 4, pp. 279–294, Dec. 2024, [Online] Available: Springer Nature Link https://link.springer.com/article/10.1007/s44366-024-0031-x [Accessed Mar. 04, 2025.], https://doi.org/10.1007/s44366-024-0031-x

[19] C. N. Akpen, S. Asaolu, S. Atobatele, H. Okagbue, and S. Sampson, "Impact of online learning on student's performance and engagement: a systematic review," Discov Educ, vol. 3, no. 1, p. 205, Nov. 2024, [Online] Available: Springer Nature Link https://link.springer.com/article/10.1007/s44217-024-00253-0 [Accessed Mar. 04, 2025.], https://doi.org/10.1007/s44217-024-00253-0

[20] L. Yuerong, M. Na, Y. Xiaolu, and S. S. Alam, "Self-determination and perceived learning in online learning communities," Sci Rep, vol. 14, no. 1, p. 24538, Oct. 2024, [Online] Available: Nature https://www.nature.com/articles/s41598-024-74878-4 [Accessed Mar. 04, 2025.], https://doi.org/10.1038/s41598-024-74878-4

[21] M. A. Al Mamun and G. Lawrie, "Cognitive presence in learner–content interaction process: The role of scaffolding in online self-regulated learning environments," J. Comput. Educ., vol. 11, no. 3, pp. 791–821, Sep. 2024, [Online] Available: Springer Nature Link https://link.springer.com/article/10.1007/s40692-023-00279-7 [Accessed Mar. 04, 2025.], https://doi.org/10.1007/s40692-023-00279-7

[22] Z. Li et al., "Students' online learning adaptability and their continuous usage intention across different disciplines," Humanit Soc Sci Commun, vol. 10, no. 1, pp. 1–10, Nov. 2023, [Online] Available: Nature https://www.nature.com/articles/s41599-023-02376-5 [Accessed Mar. 04, 2025.], https://doi.org/10.1057/s41599-023-02376-5

[23] D. Delen, Predictive Analytics: Data Mining, Machine Learning and Data Science for Practitioners, 2nd ed. 2020. [Online]. Available: https://www.oreilly.com/library/view/predictive-analytics-data/9780135946527/ [Accessed: Mar. 04, 2025]

[24] "Moodle 4.3 Database." [Online]. Available: https://www.examulator.com/er/4.3/ [Accessed: Mar. 04, 2025.]

[25] "Moodle2erd.png (3854×2641)." [Online]. Available: https://docs.moodle.org/dev/images_dev/5/5a/Moodle2erd.png [Accessed: Mar. 04, 2025.]

[26] "Large installations - MoodleDocs.". [Online]. Available: https://docs.moodle.org/405/en/Large_installations [Accessed: Mar. 05, 2025]

[27] "Events API - MoodleDocs." [Online]. Available: https://docs.moodle.org/dev/Events_API [Accessed: Mar. 04, 2025.]

[28] J. Han, M. Kamber, and Pei, Data Mining: Concepts and Techniques, 3rd ed. Morgan Kaufmann, 2011. [Online]. Available: https://www.oreilly.com/library/view/data-mining-concepts/9780123814791/ [Accessed: Mar. 04, 2025.]

[29] S. Jurenoka, "Development of Methods and Models for Generating the Adaptive Curriculum Based on the Student Knowledge Level," PhD, Riga Technical University, Latvia, Riga, 2023. [Online] Available: RTU Press https://ebooks.rtu.lv/product/development-of-methods-and-models-for-generating-the-adaptive-curriculum-based-on-the-student-knowledge-level/?lang=en [Accessed Mar. 04, 2025.], https://doi.org/10.7250/9789934229541

[30] S. Jurenoka and J. Grundspenkis, "Development of Methods and Models for Generating an Adaptive Learning Plan Based on the User's Level of Knowledge," BJMC, vol. 11, no. 1, 2023, [Online] Available: Baltic J. Modern Computing https://www.bjmc.lu.lv/fileadmin/user_upload/lu_portal/projekti/bjmc/Contents/11_1_06_Jurenoka.pdf [Accessed Mar. 04, 2025.], https://doi.org/10.22364/bjmc.2023.11.1.06