

Use of AI machine learning algorithms to assess medical student engagement and predict performance

Samar Nagah El-Beshbishi^{1,2*}, Mohammed Abdel Razek³, Hala M. El-Marsafawy^{4,5}, and

Omayma Hamed⁶

¹MSc, PhD, Professor & Head of Department of Medical Parasitology, Member at Department of Medical Education, Faculty of Medicine, Mansoura University, Mansoura, Egypt

²Department of Basic Medical Science, Faculty of Medicine, New Mansoura University, New Mansoura, Egypt.

³MSc, PhD, Professor, Department of Mathematics and Computer Science, Faculty of Science, Al-Azhar University, Cairo, Egypt

⁴Professor of Artificial Intelligence, Armed Force Faculty of Medicine (AFCM), Cairo, Egypt

⁵M.Sc., MD, Professor, Department of Pediatrics Cardiology, Faculty of Medicine, Mansoura University, Mansoura, Egypt.

⁶Department of Clinical Medical Sciences, Professor & Dean of Faculty of Medicine, New Mansoura University, New Mansoura, Egypt.

⁷M.Sc., PhD, Professor & Director of Medical Education, Medical Education Administration, Armed Forces College of Medicine, Cairo, Egypt.

Abstract

Background: To understand how students engage with course activities, it is important for educators to predict students' degree of participation. Artificial intelligence (AI) has become a valuable tool in higher education, particularly in predicting students' academic engagement. This study compares nine AI machine learning algorithms to determine students' engagement in a basic medical science course and examine its correlation with their assessment scores. Altair RapidMiner studio software was employed for data visualization, calculation of correlation coefficient, and predictive analysis. **Methods:** We employed machine learning (ML) classification algorithms to analyze students' engagement in a Medical Parasitology course taught to first year medical students. The independent variables used included their performance scores, and their level of interaction with course materials on the learning management system, such as frequency of viewing content and completing activity. The dependent variable was students' engagement levels across various activities. To predict students' engagement, we applied nine ML algorithms to the dataset: namely Naïve Bayes,

Generalized Linear Model, Logistic Regression, Fast Large Margin, Deep Learning, Decision Tree, Random Forest, Gradient Boosted Tree, and Support Vector Machine. Their performance was evaluated using several metrics. **Results:** The Logistic Regression exhibited the highest performance among the models tested, achieving an accuracy of 95%, classification error of 5%, precision of 100%, recall of 88.4%, F-measure of 93.8%, sensitivity of 88.4%, specificity of 100%. **Discussion and Conclusions:** The number of student logins to course materials was strongly related to students' engagement. Highly engaged students achieved better results on assessments compared to those with lower engagement. Additionally, students with minimal engagement participated less frequently in various course activities. These findings highlight the potential use of RapidMiner as an effective AI tool for educational institutions to accurately classify students as engaged or non-engaged.

Keywords: RapidMiner, artificial intelligence, predict, student engagement, machine learning, Logistic Regression

Date submitted: 29-January-2025

Email: Samar Nagah El-Beshbishi (selbeshbishi@mans.edu.eg)

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

Background

Student engagement in learning activities is considered a primary predictor of effective electronic learning and student performance.¹⁻³ Online learning platforms provide an enormous amount of data about student actions, such as use of

Citation: Nagah El-Beshbishi S, Abdel Razek M, El-Marsafawy H, and Hamed O. Use of AI machine learning algorithms to assess medical student engagement and predict performance. *Educ Health* 2025;38:132-144

Online access: www.educationforhealthjournal.org

DOI: 10.62694/efh.2025.272

Published by The Network: Towards Unity for Health

reading materials, videos, and quizzes,⁴⁻⁶ with each student having a unique approach to receiving and analyzing information.⁷ There is a variety of machine learning (ML) algorithms such as: Naïve Bayes, Generalized Linear Model, Logistic Regression, Fast Large Margin, Deep Learning,

Decision Tree, Random Forest, Gradient Boosted Tree, and Support Vector Machine, that could be used to predict students' engagement by identifying patterns in student behavior to predict future performance.

One of the artificial intelligence (AI) tools that uses this wide variety of ML algorithms to help in data visualization, calculation of correlation coefficient, and predictive analysis is Altair RapidMiner studio software.⁸ Intelligent predictive analytics systems analyze educational data stored in learning management system (LMS) logs and can predict the degree of student engagement.⁹ Predictive models can help teachers detect low- or non-engaged students based on their course activities. This may enable timely intervention to address difficulties, motivate students, and enhance their academic performance and engagement.^{10,11} Thus, predicting student engagement may help improve both teaching and learning processes.¹²

Researchers have used various AI techniques to evaluate students' engagement in e-learning.¹³ Most studies have used machine learning (ML) to predict at-risk students, employing classification as the data mining technique for subsequent analysis.¹⁴ Ayouni et al.¹¹ at the College of Computer and Information Science, Princess Nourah Bint Abdul Rahman University, Saudi Arabia, revealed that the intelligent predictive system could alert teachers when student engagement decreases.⁴ In Egypt, Halawa et al.¹⁵ developed a model using learners' engagement data from the social network and LMS at Business College, German University. The model helped students to be aware of their personalities, and helped teachers to match students to their learning styles.

Most related studies have relied on one algorithm, which affects results accuracy. Additionally, research on predicting students' engagement has primarily focused on intermediate and final-year students, despite the importance of understanding first-year students' experiences for early intervention.¹⁶ This study compares the performance of nine machine learning algorithms used by RapidMiner AI software to predict first-year medical student engagement and correlate it with their performance outcomes.

Methods

Study setting

This pilot study was conducted at the Faculty of Medicine, New Mansoura University (NMU), Egypt during the 2022–2023 academic year. The

research was set in a 14-week spring course. All the materials for the Medical Parasitology course were uploaded on the Moodle LMS.

Study design

A cross-sectional study design was employed.

Study population and sampling

A total of 363 first-year undergraduate medical students aged 17–20 years who were enrolled in the Medical Parasitology course.

Inclusion criteria: A 350 medical students (212 males and 138 females) who were actively learning Medical Parasitology and provided informed consent were included.

Exclusion criteria: Thirteen students were excluded due to inactive LMS account or failure to complete the final assessment.

Data collection

Data sources

The Medical Parasitology course was hosted on the LMS of Faculty of Medicine, NMU, where students' activities were recorded. Data were extracted from the Moodle database showing frequency of viewing/visiting each activity and its completion. Student performance records were exported from the student information system of Faculty of Medicine to Excel sheets.

Data processing stage

Machine learning algorithms were applied to model the associations between input variables (activity logs and students' performance records). The collected raw data underwent cleaning, transformation and normalization for accurate engagement calculations.

Nine attributes from the dataset were utilized: student viewing or completing an activity concerned with course specifications; announcements (posts); assignment; lectures; lecture notes; practical notes; case scenarios; recorded lectures; and relevant videos. Thirteen attributes from the students' information system were identified: student name; national identification number (ID); student academic ID; email address; age; gender; admission score; assignment; quizzes; midterm; objective structured practical examination (OSPE); final scores; and grades. Four irrelevant attributes (student name; national and student academic ID; and email

address) were removed to improve prediction accuracy.

Data analysis tool

Altair RapidMiner studio software (Version 10.2), a data science platform, was employed as an AI tool for data visualization, calculation of correlation coefficient, and predictive analysis.

RapidMiner was selected for its wide range of algorithms and modeling techniques such as Naïve Bayes, Generalized Linear Model, Logistic Regression, Fast Large Margin, Deep Learning, Decision Tree, Random Forest, Gradient Boosted Tree, and Support Vector Machine that can automatically detect and visualize relationships between variables.¹⁷

Data analysis methods

Engagement per course material

Engagement score = Number of times a student accessed activity ÷ Total access by all students. Student who completed an activity received a score of 1 for that activity.

Total engagement score percentage

- a. **Total number of activities completed:**
Assignment: 1 point
Case scenarios, Course specs, Lectures, Lecture notes, Practical notes, Recorded lectures, Posts, and Videos: 1 point each
- b. **Frequency of activity:**
0–5 interactions: 1 point
6–10 interactions: 2 points
11–20 interactions: 3 points
20+ interactions: 4 points
- c. **Weighing:**
Total activities: 50%
Frequency of activity: 50%
- d. **Scoring:**
Total possible points (13) = Total activities (9 points) + Frequency of activity (4 points)
- e. **Total points per user ID:**
 Total points/student were calculated and normalized on a scale of 0–100.
Normalized score = Sum of points ÷ Total possible points × 100.

Machine learning classification model

Classification in ML predicts categories based on input data using a supervised learning approach where labeled input data trains the model to predict corresponding output.¹⁸ To implement a ML

classifier, the necessary model package was imported, and the dataset loaded.

Data preprocessing and cleaning steps were performed to handle null values, duplicates, and invalid entries. This research employed binary classification to predict student engagement as: “YES” or “NO”. Nine ML classification algorithms were trained: Naïve Bayes, Generalized Linear Model, Logistic Regression, Fast Large Margin, Deep Learning, Decision Tree, Random Forest, Gradient Boosted Trees, and Support Vector Machine.

Machine Learning Model Performance Evaluation

Model performance was evaluated using the tested dataset with metrics, including: accuracy, classification error, precision, recall, F-Measure, sensitivity and specificity, and confusion matrix.¹⁹

Metrics were calculated as follows:

$$\text{Accuracy} = (TP + TN) / (FP + FN + TP + TN)$$

$$\text{Precision} = TP / (FP + TP)$$

$$\text{Recall} = TP / (FN + TP)$$

$$\text{F-Measure} = 2 \times \text{Recall} \times \text{Precision} / (\text{Recall} + \text{Precision})$$

$$\text{Sensitivity} = [TP / (TP + FN)] \times 100$$

$$\text{Specificity} = [TN / (FP + TN)] \times 100$$

Where TP, TN, FP and FN mean true positive, true negative, false positive, and false negative, respectively.

A “confusion matrix” was used to visualize model performance in matrix form, showing correct and incorrect predictions per class and identifying classes being confused by the model.

Statistical analysis

Pre-coded data were entered into the statistical package of social science software program (SPSS), version 23 for statistical analysis. Data were summarized using mean and standard deviation (SD) for quantitative variables, and number and percent for qualitative variables.

Ethical and legal considerations

Approval of the Research Ethics Committee of the Armed Force College of Medicine (AFCM), Egypt was obtained (#179, 11/02/2023). The study adhered to the principles of the Declaration of Helsinki. All data have been anonymized and used under modified attribution.

Informed consent was obtained from all students after explanation of study details. Students

understood that participation was voluntary and would not affect their faculty enrollment or learning.

Results

The study comprised 350 students aged 17–20 years. The gender distribution was 60.6% male and 39.4% female. Analysis of admission scores for the Faculty of Medicine at New Mansoura University demonstrated a mean score of 86.2% (range: 70.91%- 96.87%).

Table 1: Academic performance scores and grade distribution of the enrolled students

Scores/Grades	Study group (n=350)
Quiz 1 score	4.6/5±0.6
Quiz 2 score	4.9/5±0.5
Assignment score	5/5±0.2
Midterm score	9.2/10±1.2
OSPE	12.3/15±3.1
Final score	17.4/20±3
Grade	No. (%)
Excellent	224 (64.0)
Very good	45 (12.9)
Good	21 (6.0)
Pass	41 (11.7)
Fail	19 (5.4)

Academic achievement was evaluated using continuous and final assessment scores, and grade distribution, which served as key indicators of student performance (Table 1).

Students’ interactions with course materials were tracked through Moodle logs, and total points were

calculated along with overall engagement score percentage. Engagement scores ranged from 5% to 100%. Statistical analysis using RapidMiner revealed no significant relationships between the dependent variable and any of the independent variables examined (age, admission score, student engagement score, and performance metrics), as all

Table 2: Academic performance scores and grade distribution of the enrolled students

	Age	Admission score	Total Eng. score	Midterm score	Quiz1 score	Quiz2 score	Assignment score	OSPE score	Final score	
Age	r	1	-0.361**	-0.021	-0.044	0.034	-0.077	-0.015	-0.041	-0.063
	p		0.000	0.697	0.409	0.523	0.150	0.774	0.441	0.237
Admission score	r	-0.361**	1	-0.040	0.000	-0.031	0.076	0.084	-0.012	0.003
	p		0.000	0.459	0.993	0.564	0.155	0.116	0.824	0.960
Total Eng. score	r	-0.021	-0.040	1	0.005	0.031	0.033	-0.033	0.030	0.201
	p	0.697	0.459		0.927	0.567	0.538	0.535	0.572	0.743
Midterm score	r	-0.044	0.000	0.005	1	0.125*	0.081	0.158**	0.648**	0.745**
	p	0.409	0.993	0.927		0.020	0.131	0.003	0.000	0.000
Quiz1 score	r	0.034	-0.031	0.031	0.125*	1	0.149**	0.204**	0.182**	0.162**
	p	0.523	0.564	0.567	0.020		0.005	0.000	0.001	0.002
Quiz2 score	r	-0.077	0.076	0.033	0.081	0.149**	1	0.458**	0.101	0.086
	p	0.150	0.155	0.538	0.131	0.005		0.000	0.059	0.107
Assignment score	r	-0.015	0.084	-0.033	0.158**	0.204**	0.458**	1	0.236**	0.173**
	p	0.774	0.116	0.535	0.003	0.000	0.000		0.000	0.001
OSPE score	r	-0.041	-0.012	0.030	0.648**	0.182**	.101	0.236**	1	0.742**
	p	0.441	0.824	0.572	0.000	0.001	0.059	0.000		0.000
Final score	r	-0.063	0.003	0.201	0.745**	0.162**	0.086	0.173**	0.742**	1
	p	0.237	0.960	0.743	0.000	0.002	0.107	0.001	0.000	

Eng.: Engagement.

* Correlation is significant.

** Correlation is highly significant, suggesting a statistically significant relationship between the given independent variable and the dependent variable.

p-values exceeded 0.05. However, a significant negative correlation was identified between student age and admission scores ($r = -0.361$, $p < 0.001$), indicating that younger students tend to achieve higher admission (Table 2).

Correlation analysis between student engagement attributes, final score, and total engagement score demonstrated moderate to very strong positive correlations (0.41–0.82) among engagement attributes, except for assignment-related engagement. Total engagement score exhibited

Table 3: Correlation analysis of engagement attributes, final scores, and total engagement score

	Class = NO	Eng. in assignment	Eng. in case scenarios	Eng. in course specs.	Eng. in lecture notes	Eng. in lectures	Eng. in posts	Eng. in practical notes	Eng. in recorded lectures	Eng. In videos	Final score	Total Eng.
Class = NO	1	-0.066	-0.381	-0.352	-0.433	-0.353	-0.406	-0.426	-0.402	-0.322	-0.187	-0.859
Eng. in assignment	-0.066	1	0.015	-0.027	0.024	0.017	0.014	0.037	-0.006	0.041	0.017	0.089
Eng. in case scenarios	-0.381	0.015	1	0.559	0.757*	0.591	0.580	0.818**	0.691*	0.559	0.190	0.395
Eng. in course specs.	-0.352	-0.027	0.559	1	0.634*	0.490	0.777*	0.600*	0.631*	0.413	0.132	0.364
Eng. in lecture notes	-0.433	0.024	0.757*	0.634*	1	0.665*	0.684*	0.821**	0.782*	0.593	0.161	0.447
Eng. in lectures	-0.353	0.017	0.591	0.490	0.665*	1	0.549	0.645*	0.569	0.459	0.125	0.324
Eng. in posts	-0.406	0.014	0.580	0.777*	0.684*	0.549	1	0.634*	0.701*	0.553	0.112	0.420
Eng. in practical notes	-0.426	0.037	0.818**	0.600*	0.821**	0.645*	0.634*	1	0.779*	0.579	0.166	0.447
Eng. in recorded lectures	-0.402	-0.006	0.691*	0.631*	0.782*	0.569	0.701*	0.779*	1	0.630*	0.124	0.451
Eng. in videos	-0.322	0.041	0.559	0.413	0.593	0.459	0.553	0.579	0.630*	1	0.086	0.350
Final score	-0.187	0.017	0.190	0.132	0.161	0.125	0.112	0.166	0.124	0.086	1	0.201
Total Eng.	-0.859	0.089	0.395	0.364	0.447	0.324	0.420	0.447	0.451	0.350	0.201	1

Eng.: Engagement.
 *Strong correlation.
 **Very strong correlation

weak to moderate positive correlations (0.32-0.45) with all other engagement attributes. Additionally, final score displayed weak positive correlations with both individual engagement factors (0.09–0.20) and total engagement (0.20), as shown in (Table 3).

RapidMiner was also employed to predict students’ engagement in the course by analyzing their interactions with various course materials through a decision tree model. This model utilized different

engagement metrics to categorize students into two distinct groups: engaged [Yes; 141 (40.3%)] and not engaged [No; 209 (59.7%)]. The Decision Tree primarily used engagement with practical notes as the key predictor of overall student engagement. (Figure 1a, b).

A comparative analysis of nine machine learning (ML) models implemented using RapidMiner was conducted to evaluate their performance metrics. Among the models, Logistic Regression achieved

the highest accuracy (95%), while Naïve Bayes had the lowest (83%). Regarding classification error, Logistic Regression exhibited the lowest error rate (5%), whereas Naïve Bayes had the highest (17%). Concerning computational efficiency, Logistic Regression and Decision Tree were the fastest, completing training and scoring in two seconds, while Gradient Boosted Trees was the slowest,

requiring 20 seconds. Precision was 100% for all models except Naïve Bayes, which achieved 79.6%. Logistic Regression demonstrated the highest recall (88.4%), while Naïve Bayes had the lowest (74.4%). Similarly, Logistic Regression achieved the highest F-measure (93.8%), compared to Naïve Bayes, which had the lowest (76.9%).

Table 4: Performance comparison of machine learning models on the sample dataset using RapidMiner

	Naïve Bayes	Generalized linear model	Logistic Regression	Fast Large Margin	Deep Learning	Decision Tree	Random Forest	Gradient Boosted Trees	Support Vector Machine
Accuracy	83%	92%	95%	90%	94%	89%	89%	89%	89%
Classification Error	17%	8%	5%	10%	6%	11%	11%	11%	11%
Total Time	7s	3s	2s	18s	3s	2s	7s	20s	3s
Precision	79.6%	100%	100%	100%	100%	100%	100%	100%	100%
Recall	74.4%	80.4%	88.4%	77.2%	84.2%	74.8%	74.8%	74.8%	74.8%
F-measure	76.4%	88.9%	93.8%	86.9%	90.7%	85.6%	85.6%	85.6%	85.6%
Sensitivity	74.4%	80.2%	88.4%	77.4%	83.2%	74.8%	74.8%	74.8%	74.8%
Specificity	88.8%	100%	100%	100%	100%	100%	100%	100%	100%

In terms of sensitivity, Logistic Regression led with 88.4%, whereas other models ranged from 74.4% to 84.2%. Specificity was 100% for all models except Naïve Bayes, which had a specificity of 88.8% (Table 4).

Analysis of the confusion matrices showed that Logistic Regression outperformed other models, achieving a precision of 91.8% (true NO) and 100% (true YES). Its recall scores were 100% (true NO) and 88.4% (true YES), making it the most reliable classification model among those assessed (Table 5).

Discussion

RapidMiner analysis demonstrated that overall student engagement across course materials did not strongly correlate with final grades. However, significant correlations were observed among different engagement metrics, except for assignments. Conversely, Kuzminykh et al.²⁰ found positive correlations between engagement and academic performance, as well as between initial subject engagement and overall engagement. This divergence in findings may be attributed to differences in engagement measurement; factors such as student learning styles, and time management. Engagement measurement methods,

socio-economic status, personality traits, age, and gender, all moderate the relationship between engagement and academic performance.²¹

Furthermore, the total engagement score did not significantly correlate with performance metrics, suggesting that our engagement metrics may not predict student grades in this course. This could indicate that our measurements fail to capture behaviors influencing performance, or that this small sample showed minimal connection between engagement and grades. More specific behavioral measures such as student attendance, participation, or time spent on coursework, and qualitative data might provide better insights into how engagement could be enhanced to support learning.²²

Our results align with findings from online learning research, where performance is influenced by multiple factors. Lu et al.²³ demonstrated that early-semester engagement patterns could predict final performance, emphasizing the importance of early intervention strategies. Several critical factors can affect students' academic performance such as prior academic preparation and skills, cognitive ability, personal factors, mental health, social support, and financial factors. A more holistic approach to engagement measurement may be necessary to fully understand its impact on learning outcomes.

Table 5: Confusion matrix for the nine models on the sample dataset

Naïve Bayes			
	True NO	True YES	Class Precision
Prediction NO	56	10	84.8%
Prediction YES	7	27	79.6%
Class recall	88.9%	74.4%	
Generalized-Linear Model			
	True NO	True YES	Class Precision
Prediction NO	59	8	88.1%
Prediction YES	0	33	100%
Class recall	100%	80.4%	
Logistic Regression			
	True NO	True YES	Class Precision
Prediction NO	56	5	91.8%
Prediction YES	0	39	100%
Class recall	100%	88.4%	
Fast Large Margin			
	True NO	True YES	Class Precision
Prediction NO	56	10	84.8%
Prediction YES	0	34	100%
Class recall	100%	77.2%	
Deep Learning			
	True NO	True YES	Class Precision
Prediction NO	62	6	91.1%
Prediction YES	0	32	100%
Class recall	100%	84.2%	
Decision Tree			
	True NO	True YES	Class Precision
Prediction NO	56	11	83.6%
Prediction YES	0	33	100%
Class recall	100%	74.8%	
Random Forest			
	True NO	True YES	Class Precision
Prediction NO	56	11	83.6%
Prediction YES	0	33	100%
Class recall	100%	74.8%	
Gradient Boosted Trees			
	True NO	True YES	Class Precision
Prediction NO	56	11	83.6%
Prediction YES	0	33	100%
Class recall	100%	74.8%	
Support Vector Machine			
	True NO	True YES	Class Precision
Prediction NO	56	11	83.6%
Prediction YES	0	33	100%
Class recall	100%	74.8%	

Student engagement represents a quantification of the level of commitment and effort they invest in activities, which contribute to their persistence and achievement of learning outcomes.²⁴ Researchers

have developed methods to evaluate learning involvement to reduce attrition. Artificial intelligence models represent a significant application in education, used to predict and monitor students' engagement,²⁵ academic

performance,²⁶ and identify at-risk students.²⁷ ML algorithms have been commonly implemented to predict students' academic performance, by processing extensive data, enhancing understanding of learning processes in online educational context.²⁷⁻²⁹

Artificial intelligence was instrumental in understanding engagement impact. RapidMiner effectively mined data, revealing a negative correlation between age and admission scores, suggesting younger students tend to achieve higher scores. However, Pacheco-Mendoza et al.³⁰ found a positive correlation between age and academic performance. Such discrepancies may be due to age-related differences in learning and adaptive strategies.

Alyahyan and Düstegör proposed that classification and regression analysis are effective prediction models,³¹ that utilize algorithms to estimate the values of dependent variables based on the independent variables.³² Decision trees are graphical models that represent possible decision outcomes based on specific conditions, useful for classification, regression, and prediction tasks. RapidMiner offers several advantages for constructing decision trees, including a user-friendly graphical interface; a built-in data preparation model for handling missing values and normalization, and a multi-level decision-making approach based on engagement levels, allowing it to capture non-linear relationships in data. Moreover, it generates interactive visualizations of the decision trees and performance metrics, and exports data to various formats.³³

In this study, the Decision Tree model used engagement (practical notes, assignment, etc.) to predict students' engagement classification (YES/NO) based on course interaction metrics. However, some of the leaf nodes displayed imbalanced classes (either all YES or NO) suggesting potential model overfitting. Several leaf nodes had very few samples (2–3), insufficient for robust predictions, highlighting the need for more representative data.

Our RapidMiner ML model comparisons showed that logistic regression had the best performance metrics: accuracy (95%); classification error (5%); precision (100%); recall (88.4%); F-measure (93.8%); sensitivity (88.4%); and specificity (100%). While Naive Bayes was the worst (83%, 17%, 79.6%, 74.4%, 76.4%, 74.4%, 88.8%, for each

variable respectively), other algorithms were in-between.

Comparatively, Hussain et al.³⁴ used the Open University Learning Analytics Dataset (OULAD) to predict students' engagement in a social science e-learning course. Among six classifiers (J48, JRIP, Decision Tree, Gradient-Boosting Trees, Naive-Bayes, and classification and regression tree), J48 outperformed others with an accuracy of 88.52 % and a recall of 93.4%. Raj and Renumol³⁵ predicted students' engagement in social science courses over consecutive years in Virtual Learning Environment (VLE) using OULAD data. Their Random Forest classification algorithm achieved impressive results (95% accuracy, 95% precision, and 97.4% recall), compared to our results (89%, 100%, and 74.8%, respectively).

Conversely, Orji and Vassileva³⁶ reported lower accuracy results, using activity logs to analyze students' engagement and its relation to academic performance. Employing Random Forest their findings revealed an accuracy of 84.10% (compared to 89% in our study), reinforcing that students' engagement and assessment scores are strong predictors of academic performance. Ayouni et al.¹¹ used ML algorithms, including Decision Tree, and Support Vector Machine, to predict students' engagement levels based on LMS records, reporting 80% accuracy for Support Vector and 75% for Decision Tree (compared to 89% in our study). Variations in performance metrics can be attributed to different AI tools, measured variables, and educational context.

The confusion matrix employed in this study provided deeper insight into the classification model errors, enabling thorough evaluation beyond accuracy alone. Notably, our Logistic Regression exhibited high precision and recall for both classes, indicating reliable performance in both positive and negative predictions, and no significant skew.

Comparatively, Alruwais and Zakariah³⁷ applied various ML algorithms to OULAD data to determine the best algorithm for predicting students' engagement in VLE courses. Their findings revealed CatBoost as the best model, achieving an accuracy of 92.23%, a precision of 94.40%, and a recall of 100%. Their normalized confusion matrix revealed that 94% of negative classes were correctly predicted along with 87% of positive classes. The authors demonstrated that recall is the primary metric for identifying low-engaged students, while

accuracy is key for predicting high-engaged students.

Additionally, Alvi et al.³⁸ investigated factors affecting undergraduate Computer Science students' academic performance at IQRA University, Karachi, including social media impact on future achievements. Using RapidMiner software and the Naive Bayesian algorithm, their findings showed progressively improving accuracy across the three dataset folds (85.0%, 90.0%, and 92.7%), indicating an increasing effectiveness of the model in predicting students' academic performance.

Limitations

The study has several limitations, including a relatively small sample size (350 students), gender imbalance (212 males vs. 138 females), and its focus on a single basic medical science course at one University. Additionally, the application of a purely quantitative approach can't capture students' perceptions of online engagement. Therefore, future mixed-methods research should consider a multi-institutional study with a larger sample size and balanced gender representation, assessing both basic and clinical sciences, before generalization of results. Also, taking a qualitative approach would add to the knowledge base of the subject,

Conclusion

AI-driven learning analytics can analyze large amounts of data from the online learning environment. RapidMiner has proven to be an effective, user-friendly AI tool for classifying students based on their engagement levels, enabling

educators to identify low-engaged students who require early intervention, thereby improving overall educational outcomes. Our preliminary study revealed that Logistic Regression is the best model for predicting students' engagement. Future research should incorporate behavioral metrics and students' perceptions to better understand how increased engagement influences academic performance.

Acknowledgment

We would like to express our greatest gratitude to Prof. Meawad Mohammed ElKholy, President of New Mansoura University (NMU), New Mansoura, Egypt for his support and encouragement, and giving us the permission to use students logs and reports on LMS of Faculty of Medicine. We would also wish to thank all staff members who support us in collecting data. Finally, we would like to thank all participants for their contribution in the success of our work.

Authors' contributions

SNB and MAR conceived and designed the study proposal; SNB conducted the study as part of her Master of Health Profession Education (MHPE) thesis, reviewed the literature, analyzed the results, and wrote the manuscript; MAR helped in data curation, analysis and interpretation, software, and editing the manuscript; HMM helped in data acquisition and study supervision, and participated in reviewing the manuscript; and OH was the main supervisor of the study and helped in writing the manuscript and its final review for publication.

References

1. Ting JKS, Tan VM, Voon ML. Factors influencing student engagement in higher education institutions: Central to sustainability and progression of the institution. *European Journal of Molecular and Clinical Medicine*. 2020; 7(8): 473-486. Available at: https://ejmcm.com/article_3175.html.
2. Lee JS. The relationship between student engagement and academic performance: Is it a myth or reality? *The Journal of Educational Research*. 2014; 107(3): 177-185. <https://doi.org/10.1080/00220671.2013.807491>.
3. Hlaoui Y, Hajjej F, Ayed L. Learning analytics for the development of adapted e-assessment workflow system: CLOUD-AWAS. *Computer Applications in Engineering Education*. 2016; 24(6): 951-966. <https://doi.org/10.1002/cae.21770>.
4. Al-Fraihat D, Joy M, Masa'deh R, Sinclair J. Evaluating e-learning systems success: An empirical study.

Computers in Human Behavior. 2020; 102(1): 67-86. <https://doi.org/10.1016/j.chb.2019.08.004>.

5. Jiang P, Wang X. Preference cognitive diagnosis for student performance prediction. *IEEE Access*. 2020; 8: 219775–219787. DOI: <https://doi.org/10.1109/ACCESS.2020.3042775>.

6. Albreiki B, Zaki N, Alashwal H. A systematic literature review of student' performance prediction using machine learning techniques. *Education Sciences*. 2021; 11(9): 552. DOI: <https://doi.org/10.3390/educsci11090552>.

7. Bradley VM. Learning management system (LMS) use with online instruction. *International Journal of Technology in Education*. 2021; 4(1): 68-92. DOI: <https://doi.org/10.46328/ijte.36>.

8. Chen X, Zou D, Xie H, Cheng G, Liu C. Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology and Society*. 2022; 25(1): 28-47. Available at: <https://www.jstor.org/stable/48647028>

9. Tempelaar DT, Rienties B, Giesbers B. In search for the most informative data for feedback generation: learning analytics in a data-rich context. *Computers in Human Behaviour*. 2015; 47: 157-167. DOI: <https://doi.org/10.1016/j.chb.2014.05.038>.

10. Dixon MD. Measuring student engagement in the online course: The online student engagement scale (OSE). *Online Learning Journal*. 2015; 19(4). DOI: <https://doi.org/10.24059/olj.v19i4.561>.

11. Ayouni S, Hajje F, Maddeh M, Al-Otaibi S. A new ML-based approach to enhance student engagement in online environment. *PLoS ONE*. 2021; 16(11): e0258788. DOI: <https://doi.org/10.1371/journal.pone.0258788>.

12. Kuzilek J, Hlosta M, Zdrahal Z. Open University Learning Analytics dataset. *Scientific Data*. 2017; 4: 170171. DOI: <https://doi.org/10.1038/sdata.2017.171>.

13. Costa EB, Fonseca B, Santana MA, de Araújo FF, Rego J. Evaluating the effectiveness of educational data mining techniques for early prediction of student academic failure in introductory programming courses. *Computers in Human Behavior*. 2017; 73: 247–256. DOI: <https://doi.org/10.1016/j.chb.2017.01.047>.

14. Akçapınar G, Altun A, Aşkar P. Using learning analytics to develop early-warning system for at-risk students. *International Journal of Educational Technology in Higher Education*. 2019; 16: 40. DOI: <https://doi.org/10.1186/s41239-019-0172-z>.

15. Halawa MS, Shehab ME, Hamed EMR. Predicting student personality based on a data-driven model from student behavior on LMS and social networks. *The Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*. October 7-9, Sierre; 2015. Available at: ieeexplore.ieee.org.

16. Chou CL, Kalet A, Costa MJ, Cleland J, Winston K. Guidelines: The dos, don'ts and don't knows of remediation in medical education. *Perspectives on Medical Education*. 2019; 8(6): 322-338. DOI: <https://doi.org/10.1007/s40037-019-00544-5>.

17. Al-Ma'aitah MA. Utilizing of Big Data and predictive analytics capability in crisis management. *Journal of Computer Science*. 2020; 16(3): 295-304. DOI: <https://doi.org/10.3844/jcssp.2020.295.304>.

18. Suhaimi NM, Abdul-Rahman S, Mutalib S, Abdul Hamid N, Hamid A. Review on predicting students' graduation time using machine learning algorithms. *International Journal of Modern Education and Computer Science*. 2019; 7: 1-13. DOI: <https://doi.org/10.5815/ijmecs.2019.07.01>.

19. Rimadana MR, Kusumawardani SS, Santosa PI, Erwianda MSF. Predicting student academic performance using machine learning and time management skill data. *International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*. Yogyakarta, Indonesia; 2019. DOI: 10.1109/ISRITI48646.2019.9034585.
20. Kuzminykh I, Ghita B, Xiao H. The relationship between student engagement and academic performance in online education. In *2021 5th International Conference on E-Society, E-Education and E-Technology (ICSET 2021)*. Association for Computing Machinery, New York, NY, USA; 2021. p. 97-101. DOI: <https://doi.org/10.1145/3485768.3485796>.
21. Francis P, Broughan C, Foster C, Wilson C. Thinking critically about learning analytics, student outcomes, and equity of attainment. *Assessment and Evaluation in Higher Education*. 2020; 5(6): 811-821. DOI: <https://dx.doi.org/10.1080/02602938.2019.1691975>.
22. Bijkerk LE, Oenema A, Geschwind N, Spigt M. Measuring engagement with mental health and behavior change interventions: an integrative review of methods and instruments. *International Journal of Behavioral Medicine*. 2023; 30: 155–166. DOI: <https://doi.org/10.1007/s12529-022-10086-6>.
23. Lu O, Huang AYQ, Huang JCH, Lin AJQ, Ogata H, Yang S. Applying learning analytics for the early prediction of students' academic performance in blended learning. *Educational Technology and Society*. 2018; 21: 220-232.
24. Sugden N, Brunton R, MacDonald J, Yeo M, Hicks B. Evaluating student engagement and deep learning in interactive online psychology learning activities. *Australasian Journal of Educational Technology*. 2021; 37: 45-65. DOI: <https://doi.org/10.14742/ajet.6632>.
25. Chen X, Xie H, Zou D, Hwang GJ. Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*. 2020; 1(3): 100002. DOI: <https://doi.org/10.1016/j.caeai.2020.100002>.
26. Mustapha MF, Zulkifli AN, Kairan O, Zizi NN, Yahya NM, Mohamad NM. The prediction of student's academic performance using RapidMiner. *Indonesian Journal of Electrical Engineering and Computer Science*. 2023; 32(1): 363-371. DOI: <http://doi.org/10.11591/ijeecs.v32.i1.pp363-371>.
27. Ouyang F, Wu M, Zheng L, Zhang L, Jiao P. Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education*. 2023; 20(1):4. DOI: <https://doi.org/10.1186/s41239-022-00372-4>.
28. Jiao P, Ouyang F, Zhang Q, Alavi AH. Artificial intelligence-enabled prediction model of student academic performance in online engineering education. *Artificial Intelligence Review*. 2022; 55(129): 6321-6344. DOI: <https://doi.org/10.1007/s10462-022-10155-y>.
29. Crompton H, Burke D. A systematic review of artificial intelligence in higher education. *International Journal of Educational Technology in Higher Education*. 2023; 20(1): 24. DOI: <https://doi.org/10.1186/s41239-023-00392-8>.
30. Pacheco-Mendoza S, Guevara C, Mayorga-Albán A, Fernández-Escobar J. Artificial intelligence in higher education: A predictive model for academic performance. *Education Sciences*. 2023;13: 990. DOI: <https://doi.org/10.3390/educsci13100990>.
31. Alyahyan E, Düstegör D. Predicting academic success in higher education: literature review and best practices. *International Journal of Educational Technology in Higher Education*. 2020; 17:3. DOI: <https://doi.org/10.1186/s41239-020-0177-7>.

32. Bramer M. Measuring the performance of a classifier. In: *Principles of Data Mining*. Mackie E, editor. London, Springer; 2016. p. 178-179.
33. Chahal H, Gulia P. Experimental evaluation of open-source data mining tools: R, Rapid Miner and KNIME. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*. 2019; 9 (1): 4133–4144. DOI: 10.35940/ijitee.A5341.119119.
34. Hussain M, Zhu W, Zhang W, Abidi SMR. Student engagement predictions in an e-learning system and their impact on student course assessment scores. *Computational Intelligence and Neuroscience*. 2018; 6347186(6): 1-21. DOI: <https://doi.org/10.1155/2018/6347186>.
35. Raj NS, Renumol VG. Early prediction of student engagement in virtual learning environments using machine learning techniques. *E-Learning and Digital Media*. 2022; 19(6): 537-554. DOI: <https://doi.org/10.1177/20427530221108027>.
36. Orji F, Vassileva J. Using machine learning to explore the relation between student engagement and student Performance. *The 24th International Conference Information Visualisation (IV)*. Melbourne, Australia; 2020. DOI: 10.1109/IV51561.2020.00083.
37. Alruwais N, Zakariah M. Student-engagement detection in classroom using machine learning algorithm. *Electronics*. 2023; 12(3): 731. DOI: <https://doi.org/10.3390/electronics12030731>.
38. Alvi I, Kazimi AB , Alvi MA, Swift S, Ahmed S. Educational Data Mining for Predicting Students' Academic Performance. *Kurdish Studies*. 2024; 12(4): 820-835. DOI: <https://doi.org/10.53555/ks.v12i4.3063>.