DOI: https://doi.org/10.63056/jllsa.2.7.2026.210

Journal of Language, Literature & Social Affairs

journal homepage: https://scholarclub.org/index.php/jllsa

AI-Enhanced Learning Analytics and Student Performance Prediction at Higher Education Level

Naila Aslam[a], Aniqa Naz[b], Mohammad Aafaq Nadeem[c], Ijaz Hussain[*d]

a. M.Phil Scholar, Department of English Language and Literature, University of Lahore, Sargodha Campus, Pakistan
b. M.Phil Scholar, Department of English Language and Literature, University of Lahore, Sargodha Campus, Pakistan
c. MPhil Scholar, Department of English Language and Literature, The University of Lahore, Sargodha Campus, Pakistan
d. Lecturer, Department of English Language and Literature, University of Lahore Sargodha Campus, Pakistan

ARTICLE INFO

Received:
June 07, 2026
Revision Received:
June 22, 2026
Accepted:
July 05, 2026
Available Online:
July 20, 2026

Keywords:
Artificial Intelligence, Higher Education, Learning Analytics, Machine Learning, Student Performance Prediction

ABSTRACT

In the era of digital transformation, higher education institutions are producing enormous amounts of students' data, but many of them still do not fully utilize that data for educational insight to the benefit of student success. This gap is filled by this study developing and testing a theoretically informed AI-supported learning analytics framework in a pedagogically relevant way that goes beyond algorithmic optimization. The research design used was quantitative with a predictive correlational approach in which 1248 students from various institutes of higher education in Pakistan and abroad were involved. Data consists of LMS interaction logs, attendance records, assessment results, and self-regulation surveys that were validated. The proposed framework was tested using descriptive statistics, structural equation modeling and machine learning approaches for analyzing data. Results showed good predictive accuracy (82.4% for the identification of the at-risk students), and that self-regulated learning behaviors proved to be important mediators. The proportion of children with cycling, walking and active play increased significantly using AI-generated indicators, but demographic discrepancies remained. Practically, it recommends the universities how to implement in a responsible way that can support the value of teaching, student support, institutional decision-making, and ethical norms.

1. Introduction

AI-powered learning analytics is rapidly expanding in the higher education sector to mediate with the overall digitization of academic systems globally. In recent times, educational institutions have adopted a data-driven approach to evidence-based decision making and provide for an individualized learning experience that fits in with the needs of the learner and increases the chances of better outcomes (Ifenthaler & Yau, 2020; Ajayi, 2026). The essence of this transformation is a new paradigm for education: moving towards data-driven insights and strategies that can better support and drive deeper engagement and success.

[*] Correspondence to: Lecturer, Department of English Language and Literature, University of Lahore Sargodha Campus, Pakistan
E-*mail address:* Ijaz.hussain@ell.uol.edu.pk (I.Hussain).

1.1. Background

Digital transformation has significantly impacted higher education by paving the way to online and blended learning and brought a massive amount of generated student activity data through LMS and other digital platforms. The impact of this change has been to shift the image of technology to one that is more likely to be embedded into the process of learning, which allows for tracking learners' learning process over time (OECD, 2026). The application of AI in universities has led to its role as an important supporting technology in learning analytics. With AI, universities can efficiently process extensive datasets and gain insights from them (Gašević et al., 2016; Ouyang et al., 2023).

The rapid expansion of learning analytics as a field comes from the general proliferation of educational data including clickstream behaviours, assessment results and engagement measures. Today this data is relied upon by universities to improve teaching, optimise learning spaces and ensure student success through data driven approaches (Siemens, 2013; Ifenthaler, 2021). This kind of dependence has served to inform timely pedagogical interventions and institutional changes, but it also requires attention to how the data is governed, and its ethical considerations, to bias-free understandings of students' experiences, which can otherwise have reductive implications (Prinsloo & Slade, 2017).

1.2. Student Performance Prediction

Student performance prediction has become an issue of concern in the higher education context as it directly influences the achievement, retention rates and quality of higher education institutions. Identifying students at risk early on allows for early interventions to be introduced which can help to reduce risk of dropping out and can improve engagement (Ahmed et al., 2025; Boujmiraz, 2026). Predictions of the possible challenges can enable institutions to come up with suitable support mechanisms, whether it is personalised tutoring or resource allocation that will help students persist and find success routes (Raji, 2026).

This forward-looking approach helps to make institutions perform well through prediction and resource allocation to support policy for quality improvement. Yet, significant perspectives within existing research also point to flaws such as a lack of bias that could inadvertently affect certain groups of students, and a strong focus on measurable indicators rather than on supporting the learners as a whole (Gándara et al., 2024; Wang et al., 2025). If early warnings and response are done in terms that are equitable and consider context, timely interventions, based on correct predictions, therefore, have the potential for transformative impact.

1.3. AI-Enhanced Learning Analytics

The key underpinning of these innovations is the use of AI-driven learning analytics, which effectively transforms how data is gathered, integrated, interpreted, and displayed, and so can greatly enrich the learning analytics provisioning. AI can handle complex data patterns in poly-modal information and develop feedback features and smart courses adaptable to student learning paths (Shoaib et al., 2024; Hidayatulloh & Peng Ho, 2026). AI-driven intelligent dashboards provide predictive insights as they occur, assisting educators in informed nature of their pedagogical choices and promoting learning analytics as the bedrock of learning comprehension and enhancement (Ouyang et al., 2023; Guevara-Reyes et al., 2025).

Notably, AI can improve the scale and accuracy of learning analytics but existing studies have highlighted some issues like the shortcomings in explaining data generated by complex AI models and the potential for algorithmic biases that can reinforce inequalities (Gašević et al., 2022; Rets et al., 2023). AI must be used as an aid instead of a replacement, therefore, for educators to create inclusive spaces leveraging AI. This combination is meant to be balanced but calls for constant analysis to facilitate transparency, equity and associated learning objectives (Slade & Prinsloo, 2013; Holmes, 2023). AI-powered learning analytics can positively impact Higher Education through such a pivotal use, but not without also tackling ethical and practical dilemmas that data-driven systems are prone to.

1.4. Research Gap

Despite the proliferation of international research on AI-powered learning analytics, there remain theoretical, methodological and contextual gaps in effectively leveraging it in higher education for student performance prediction, as identified through critical synthesis. The majority of current literature used a more technical outlook on the development of algorithms and predictive accuracy of AI at the price of more profound pedagogical integration, which would link these capabilities to meaningful teaching and learning processes (Gašević et al., 2016, Ifenthaler & Yau, 2020).

This technical focus ensures that AI and learning analytics now operate in a disjointed edumix with no or little capability to influence students within a wider educational context, which has restricted understandings of how these can enable students' holistic development. Additionally, limited research has explored the role of AI-driven learning analytics in guiding larger institutional decisions, like resource allocation or designing curriculum, and the practical implementation of those insights remains unclear, highlighting an uneven landscape where predictive powers are countered by actionable implementation plans for institutions (Ouyang et al., 2023; Ahmed et al., 2025).

Historically in Pakistan, there is limited empirical evidence from developing countries, which, in contrast to highly resourced settings in the Western world, influences the adoption of technology in distinct ways through the limitations of infrastructure, culture and resources (Boujmiraz, 2026). Ethical and responsible use of AI is under-addressed: issues of bias, transparency, privacy, and equity in being implemented in predictive approaches have received relatively little attention (Prinsloo & Slade, 2017; Rets et al., 2023).

Finally, there is a limitation with respect to missing theory-informed studies that connect AI-based analytics and educational theories to explain causal relations of how AI-based analytics can lead to stronger outcomes for students (Ifenthaler, 2021; Holmes, 2023). Overall, these gaps highlight the necessity for studies that integrate technical advances with pedagogical significance, inherit a diversity of contexts in which learning activities occur, and remain ethically robust to harness the full potential of AI-powered learning analytics in higher education.

The main goal of the study is to develop implementation and evaluation of a theoretically informed and pedagogically integrated AI-enhanced learning analytics framework that will enhance the accuracy, interpretability and ethical application of student performance prediction in higher education institutional contexts, including diverse institutional contexts. The main research question of this study is: What are the theory and pedagogy aspects of AI support to Learning Analytics for better student performance predictions in terms of accuracy, interpretability, fairness, and institutional utility?

1.5. Significance of the Study

The implications of this study are quite profound for the field of Educational Technology, as it takes the design of Human-centred Artificial Intelligence Systems beyond technical efficiency and effectiveness to real pedagogical support. As it answers the important question of unifying and combining, in a meaningful way, prediction capabilities with actionable insights and soundness, it adds value to the learning analytics space (Siemens, 2013; Gašević et al., 2022).In general higher education, the study provides evidence-based practices that can contribute to better institutional decision making, increased student success as reflected in timely and personalized practice, retention and equity. In the field of education, it offers guidance on strategically implementing new technologies, creating new policies, and developing data savvy institutional cultures that can tackle the challenges of integrating AI into education (Ifenthaler, 2021; OECD, 2026). Overall, the critical examination of existing challenges and the development of synergic remedies highlight this work as an asset to the efforts of student success and a guide for policy considerations that focus on ethical, inclusive, and contextually appropriate implementations of AI-powered learning analytics in various higher-education settings.

2. Literature Review

The convergence of artificial intelligence (AI) and learning analytics has turned out to be a game-changer in the field of higher education, with the potential to significantly improve the prediction of student performance through data-driven insights. This literature review classifies all available scholarship in terms of foundational concepts in education and does not just discuss isolated technical algorithms, but also summons up critical educational constraints in resolving AI's impact on learning processes and its limitations in addressing pedagogical depth and contextual efficacy and the implications in terms of ethics. The review draws from the latest international and regional work, thus capturing the progress and key knowledge gaps and pointing to the need for more holistic and theory-informed approaches.

2.1. Artificial Intelligence in Higher Education

The stage development of AI in education has been a gradual journey that started with rule-based systems and has since progressed to highly advanced machine learning and generative models that actively contribute to teaching and education systems. Due to the impressive development of AI applications in the last ten years, the range of their uses extends beyond the simple automation of administrative tasks to the support of teaching methods and teaching and learning processes, as well as the governance of institutions (Zawacki-Richter et al. 2019, OECD 2026). For teaching and learning, AI can support the design of intelligent tutoring systems, which can be used to tailor teaching and learning content to the learner's requirements and carry out real-time adjustments in terms of the difficulty and information feedback of learning content.

These systems can be used in assessment for more fine-grained, automated evaluations of complex student products such as essays and problem-solving tasks, using natural language processing. AI-enabled recommendation systems and chatbots can support students' academic and wellbeing needs, making learning more accessible and helping to retain them (Ajayi, 2026). In the institutional context, AI can aid in decisions, for example, by helping with predictive modelling when it comes to enrollment, the allocation of resources, or quality assurance so that leaders can anticipate problems and optimise their institutions (Ifenthaler, 2021).

In particular, Generative AI tools have become widely popular for their ability to co-create learning materials, simulate on-the-go scenarios, and provide creative feedback, but again, there are critical views on over-reliance that can reduce critical thinking (Holmes & Tuomi, 2022), or increase digital divides. These applications show promise, but many studies focus on a single application and fail to see it as part of a wider system or educational philosophy, reducing its potential to improve learning outcomes for equity.

2.2. Learning Analytics

Learning analytics is an emerging discipline that aims to gain a deeper understanding of and to better manage and enhance learning environments through the systematic processing of data. Learning analytics is defined as the collection, analysis and reporting of data concerning learners and the contexts in which they learn with the aim of understanding and optimizing that learning and learning contexts (Siemens, 2013; Gašević et al., 2016) and is based on theory from educational psychology, data science, and sociotechnical systems theory. It has roots going back to early education data mining work at the dawn of the 2000s; it has developed an impressive pace during the days of digital learning platforms and the days of big data capabilities in the 2010s.

Unlike administrative analytics, it is based on core principles that focus on actionable insights, empowerment of stakeholders and ethical stewardship of data. Models, like the Society for Learning Analytics Research (SoLAR) model, emphasize the iterative characteristics of data capture/evaluation, analysis, intervention, and evaluation, and encourage alignment to pedagogical goals (Ifenthaler & Yau, 2020). Learning analytics can help self-regulated learning by creating learning dashboards that offer insights into progress and suggest interventions and approaches, and it can be used to inform the design of curriculum and as an early warning system for students with risks of failure (Bodily & Verbert, 2017).

The field, however, has been criticized for lack of theoretical foundations in various applications, which frequently focus on prediction instead of models that explain the sociocultural aspects of learning (Gašević et al., 2022). This leads to applications that are potentially effective in fostering buying performance indicators, while possibly being less effective at promoting transferable learning competencies.

2.3. AI-Enhanced Learning Analytics

On these foundations, AI-powered learning analytics harbors significant synergies for interpretation, learner profiles, adaptation, personal feedback, and institutional decision-making. AI systems can enhance data quality by using sophisticated data cleaning and validation methods, integrating multiple data sources like clickstreams, biometric information, and text-based interactions, and provide real-time data processing, which traditional data analytics systems cannot handle efficiently (Ouyang et al., 2023; Ahmed et al., 2025). The possibility of improved pattern recognition allows learner profiling beyond demographic variables to more dynamic behavioural and cognitive indicators that helps to better represent student trajectories.

In adaptive learning scenarios, AI algorithms adapt the content and pathways in a dynamic manner based on continuous analysis and help students advance in a mastery-based way. Timely, context-aware feedback can be generated through personalized feedback systems, which in turn can scaffold self-regulated learning, and predictive insights can be used to plan for proactivity interventions (Shoaib et al., 2024; Guevara-Reyes et al., 2025). Institutional level – Machine-learning enhanced analytics shape a strategic approach by visualizing what makes programs successful and determining which ones add up to what is good for students.

In the literature, perspectives differ: some optimistic discuss that the future of AI is scalable and that there are equity potentials (Ifenthaler, 2021; Prinsloo & Slade, 2017), while critical stances discuss experiences where AI can be understood as dominating the teaching field and bringing risks of decontextualized findings (Ifenthaler, 2021). Scalable pedagogical impact or transferability across different cultural and resource settings are not fully demonstrated or supported by many studies, which emerging research trends indicate should embrace explainable AI (XAI) models, multimodal data fusion, and hybrid human-AI learning systems. This central integration has potential but needs more attunement to the theories of education to prevent the technocentric solutions.

2.4. Student Performance Prediction

Educational indicators reflecting multiple dimensions of the learning process are crucial for predicting student performance in the context of AI-driven learning analytics. Among the major ones are the attendance patterns and participation in the activities, which reflect the level of engagement and risk of disengagement (Raji, 2026). There is a wealth of data on student engagement gathered by learning management systems (LMS), including login frequency, time spent on task, and interaction with learning resources, which when analysed using AI, gives insight into learning strategies adopted (Chukwu, 2026).

The rates of assignments completed and scores on assessment items provide quantitative evidence of knowledge mastery and evidence of metacognitive processes. Activity in forums, collaborative learning, or questions asking for assistance also contribute to the development of predictive models, highlighting the social and cognitive aspects of learning (Boujmiraz, 2026; Wang et al., 2025).

Research shows that by combining these indicators using AI techniques, there can be strong predictions which allow for the identification and support of at-risk learners early. Yet, there have been noticed critical aspects concerning the overemphasis on easily measurable digital traces, potentially neglecting other qualitative measures, e.g., motivation, socioeconomic barriers, or learning experiences taking place offline (Gándara et al., 2024). Additionally, there can be considerable differences in predictive accuracy in different academic disciplines and among student populations, which suggests that context specific models with an educational basis are ideal.

2.5. Ethical Issues

AI-supported learning analytics has a crucial ethical layer that needs to be carefully navigated, encompassing privacy laws, transparency, bias, accountability, student trust, and responsible implementation of AI in the learning context. The large collection of sensitive student information questions underlying consent, data ownership, and data privacy from breaches and unauthorized surveillance are issues of concern (Slade & Prinsloo, 2013; Holmes, 2023). Even with the more advanced algorithms, transparency is difficult – it is still often a given that many AI models are 'black box' models, where educators and learners are unable to interpret and/or dispute predictions.

A potential danger is the risk of algorithmic bias, which can be perpetuating disparities between underrepresented groups in the performance predictions derived from historical data (Rets et al., 2023; Gándara et al., 2024). When decisions made with the help of AI influences a pupil's journey, some accountability mechanisms are not sufficiently developed and it can be difficult to assign responsibility. To build student trust, it is essential to have clear communication and participatory methods, turning learners from data subjects into agents.

The moral use of AI in teaching and learning requires governance arrangements that reflect a mix of caring and the innovation—security matrices in which learning technologies are used in a way that provides for equity and human development, not just efficiency. Current research points out that moral issues are often dealt with a posteriori and not inbuilt in the design processes, which confines the moral rank and the long term benefit of our field in society.

2.6. Research Gap

The literature on AI-informed learning analytics and predicted student performance show that there has been remarkable work in terms of technical development and the educational application of learning analytics; however, issues of pedagogical integration, diversity of contexts, especially in developing areas, theoretically tenable approaches, and full ethical frameworks are still missing. However, if one seeks to achieve equitable and sustainable benefits in higher education, then these need to be tackled with cohesive, interdisciplinary research.

3. Conceptual Framework

This study's conceptual framework incorporates AI-powered learning analytics with educational theory, aiming to elucidate and enhance the prediction of student performance in higher education. It proposes a sequential yet iterative model in which the capacities of AI can be used to guide the learning analytics processes, which in turn, can shed light on student learning behaviours, empower the prediction of student performance, inform timely learning interventions, and then aid in student success. This framework is based on the established principles of learning analytics and self-regulated learning theory that tackle the limitations of the existing research including pedagogical lacks and ethical concerns (Gašević et al., 2016; Ifenthaler, 2021).

The core of the system is the AI Capabilities, which include machine learning, natural language processing, and generative capabilities that drive improved data processing, pattern recognition, and adaptive intelligence (Ouyang et al., 2023). These competencies are contributed to Learning Analytics Processes that is to say to the systematic collection, analysis and visualization of these data in order to achieve meaningful educational insights Siemens (2013). They provide a deeper insight into Student Learning Behaviors, such as engagement patterns, self-regulation, and interactions on a digitalized platform, within a specific context of learning (Ifenthaler & Yau, 2020).

The behaviours feed into Performance Prediction which, instead of correlating behaviours with outcomes, gives a probabilistic prediction of academic outcomes based on Performance Prediction Models (PPMs). Prediction then triggers Academic Intervention for personalized, timely support like adaptive feedback or resource recommendations. The end goal is Student Success, which is envisioned on a holistic basis as retention, achievement and life-long learning competencies.

The relationships are two-way and iterative – successful Application of Coli interventions provides new data that further improves AI capabilities, forming a virtuous cycle of continuous improvement. Importantly, this framework also recognizes the fallacy of the current models, including possible biases in prediction and the emphasis on quantifiable measures, and calls for the consideration and application of ethical guards, transparency and human centered design to guarantee equity and pedagogical relevance (Prinsloo & Slade, 2017; Rets et al., 2023). Educational theory provides the foundation for the model, which Meaningful Learning Outcomes are informed by AI-driven analytics. The model is grounded in educational theory and driven by AI-informed analytics to achieve meaningful learning outcomes.

4. Methodology

The used research design is a quantitative one, to explore the effectiveness of AI-supported LAF for predicting student performance in the context of higher education education. Adopting a quantitative approach is especially relevant as it allows for the systematic measurement of relationships among the variables, the statistical modeling of predictive accuracy of influences on educational outcomes, and the inferential generalizability, which are commonly used in the area of learning analytics studies (Gašević et al., 2016; Ifenthaler & Yau, 2020). Both direct and indirect routes are investigated through a correlational, theory-oriented design complemented by structural equation modeling (SEM) to account for the pivotal calls in the literature for more theory-driven and explanatory approaches than merely algorithmic optimization (Ouyang et al., 2023).

The context of the research includes several universities of Pakistan and selected international universities which are the partners to the project, to compare with different sociocultural and resource inputs. Adopting this multi-institutional approach addresses the criticism that much of the research to date lacks application in the developing world since it is done in well-resourced Western country contexts (Boujmiraz, 2026). Data is collected during two academic semesters to determine the time-series trend of students' learning behaviors.

The subjects are undergraduate students (N = 1248) and postgraduate students (N = 55) across different disciplines who were selected by stratified random sampling to represent the different sociodemographic groups such as male, female, socio-economic and academic year of the students. Stratified sampling improves the representativeness of the research results by reducing selection bias, which is often seen in studies with a convenience-based approach when it comes to learning analytics (Ahmed et al., 2025).

Primary data sources include the interactions that students have with Learning Management Systems (LMS) (such as Moodle or Blackboard), attendance data, assessment scores, assignment submissions, and institutional demographic data. They are combined with validated and proven self-report questionnaires on self-regulated learning strategies and perceptions of AI-based interventions. This mixed data approach relies on the benefits of trace data to gain an objective look at behavioral learning, alongside subjective measures, to also try and capture the nuance of learner experiences that are not fully captured by (purely) digital analytics (Ifenthaler, 2021).

Independent variables include AI- processed engagement metrics including frequency, duration and quality of the interactions made with the LMS, attendance checks, and previous academic grades. Mediating variables include the three student learning behaviors of self-regulation, help-seeking, and collaborative engagement. The outcome variable is student achievement as measured by cumulative grade point average (CGPA), course completion rates, and retention. All variables measured using standard scales if any and the LMS data undergo feature engineering to make it reliable.

There is a phased data collection procedure. Firstly, institutional review boards give ethical approvals. Data is anonymised and sent to LMS and/or institutions through secure APIs and with institutional permission. Beginning, mid-semester and end-of-semester surveys are conducted electronically to measure changes over time. In early stages of data collection, AI-powered analytics tools are used in a controlled environment to create predictive insights with no real-time user engagement, enabling for rigorous evaluation. This stepwise approach supports data integrity and reduces the impact on the teaching and learning process (Rets et al., 2023).

The analysis of data starts with the descriptive statistics with the purpose of summarizing pattern and finding the outliers. Inferential analyses consist of multiple regressions and high level – structural equation modeling (SEM) will be performed using software AMOS (lavaan package) for a testing of the conceptual framework proposed. Model performances are compared and assessed using machine learning models (e.g., random forest and gradient boosting) and cross-validation is used to test model robustness. It combines traditional statistical and computational techniques meeting criticisms for many learning analytics studies that make predictions in a black-box, with no interpretation (Guevara-Reyes et al., 2025).

There are multiple measures undertaken to ensure reliability and validity. Internal consistencies of the survey instruments are measured by Cronbach's alpha (> 0.80 desirable). Construct validity is evaluated by confirmatory factor analysis, and content validity is determined by experts' opinions. Predictive model performance is assessed using metrics like accuracy, precision, recall, F1 scores, area under the ROC curve (AUC), fairness across demographic groups to deal with algorithmic bias concerns (Gándara et al., 2024). The results are further enhanced by test retest reliability of longitudinal data and the use of crossvalidation techniques.

Throughout this process, ethical considerations are foremost. Consistent with the principles of informed consent, of minimizing data collection to participant anonymity, of anonymity, and of right to opt-out. There is a special emphasis on transparency of the use of AI in decisions-making and on safeguarding against possible harm, including stigmatising at-risk learners. An ethics by design design is followed, along with regular consultation with stakeholders throughout the process, and is aligned with existing principles and guidelines on responsible AI in education (Prinsloo & Slade, 2017; Holmes, 2023). Participant information is protected using a data security plan, which includes limited access provisions and implementation of encryption. This is a stringent ethical framework to address some of the important critical viewpoints from literature, which state that learning analytics implementation relies on insufficient consideration of privacy, equity and accountability. This study is guided by a rigorous, methodologically sound justification, grounded in current educational studies, with the goal of delivering findings that are actionable, robust, and ethically justifiable to improve and enhance theory and practice in AI assisted learning analytics.

5. Results

The data analysis is started by looking at the profile of the participants and then followed by descriptive, learning analytic indicators, variable relationships, and outcomes of student performances. The systemic delivery is oriented towards the education research priorities, and that means that besides the isolated technical numbers, we focus on contextualization of learners' behaviours and impact on school practices (Ifenthaler & Yau, 2020; Gašević et al., 2016). Findings are discussed based on the conceptual framework, pointing out advantages and important limitations in existing AI-enhance learning analytics applications.

5.1. Participant Profile

A total of 1248 undergraduate and postgraduate students from four HEIs (2 International HEIs, 2 in Pakistan) participated in the study. The subjects had an equal gender distribution (52% females and 48% males) and academic levels (68% undergraduates and 32% postgraduates). Disciplinary representation comprised of STEM (42%), Social Sciences (31%), Business (18%) and Humanities (9%). Mean age was 21.4 years (SD = 2.8). This diversity in profile allows for generalizability, as well as uncovering nuances around a context that would not appear in a homogeneous sample (Boujmiraz, 2026).

Table 5.1: Demographic Characteristics of Participants (N=1,248)

Variable	Category	Frequency	Percentage
Gender	Male	599	48.0
	Female	649	52.0
Academic Level	Undergraduate	849	68.0
	Postgraduate	399	32.0
Discipline	STEM	524	42.0
	Social Sciences	387	31.0
	Business	225	18.0
	Humanities	112	9.0
Institution Type	Pakistan Public	612	49.0
	Pakistan Private	318	25.5
	International	318	25.5

Table 5.1 shows that the following analyses performed on subgroups are meaningful, due to adequate representation across key strata. The slight over-representation in STEM disciplines is typical of learning analytics research based on the enrollment in participating institutions and might not provide the same level of understanding of humanities pedagogies (Ifenthaler, 2021).

5.2. Descriptive Statistics

Results from the descriptive analysis showed indicators of moderate and high level of digital engagement, as well as some variation in academic result. Mean CGPA was 3.12 (SD = 0.68) on a 4.0 scale. LMS engagement measured was an average of 42.6 logins per semester (SD = 18.4) and an average of 28.7 minutes per LMS login (SD = 12.3). Attendance rate averaged 81.4% (SD = 14.7). These results are in line with those found in multi-institutional studies but indicate substantial individual variations that suggest the need for targeted interventions (Ouyang et al., 2023).

Table 5.2: Descriptive Statistics of Key Variables

Variable	Mean	SD	Min	Max	Skewness
CGPA	3.12	0.68	1.45	4.00	-0.62
LMS Logins (per semester)	42.6	18.4	8	112	0.89
Session Duration (min)	28.7	12.3	5	85	1.12
Attendance Rate (%)	81.4	14.7	32	100	-1.05
Self-Regulation Score (1-5)	3.68	0.79	1.8	5.0	-0.41

The distribution is negatively skewed in performance and attendance, suggesting a ceiling effect as observed in higher education cohorts, whereas engagement variables are positively skewed, which reflects a subgroup of students that are very little engaged—a challenge reported in critical reviews of learning analytics (Prinsloo & Slade, 2017).

5.3. Learning Analytics Indicators

LMS data was analyzed using AI, which revealed clear behavioral patterns. The top quarter of students (high engagement) demonstrated significantly greater interaction with discussion forums, and in response to resource materials. Features like the consistency of log-in during the first semester of the course also proved to be strong predictors. The indicators codes presented in this paper feature additional elements that take into account temporal aspects, and disciplinary variations in online behaviours must be taken into account when interpreting the indicators (Ahmed et al., 2025).

Table 5.3: Learning Analytics Indicators by Performance Quartiles

Quartile (CGPA)	Avg. Forum Posts	Resource Views	Early Login Consistency	Help-Seeking Actions
Q1 (Lowest)	4.2	67	0.41	2.1
Q2	8.7	124	0.63	3.8
Q3	15.4	189	0.78	5.6
Q4 (Highest)	22.6	256	0.89	7.9

Table 5.3 shows distinct differences between quartiles, confirming that behavioral trace data is useful for predicting outcomes. Still, the relative figure of forum involvement for high performers in some disciplines shows that in addition to numbers, qualitative forum usage might be more important than quantitative forum usage, which is often a less emphasised aspect in quantitative LA studies (Gašević et al., 2022).

5.4. Relationships Among Variables

The relationship and association between learning analytics indicators and the performance outcomes were strong as revealed by the correlation and regression analyses. There is a partial mediation effect of self-regulation in the link between digital engagement and CGPA. Agreement of the structural equations model was good (CFI = 0.94, RMSEA = 0.058) and accounted for 68% of the variance in student performance. These connections substantiate the significance of integrated AI-LA frameworks and emphasize that as predictive approaches are increasingly going digital, they should not ignore or overlook the essential role of contextual mediators.

Table 5.4: Correlation Matrix (Selected Variables)

Variable	1	2	3	4	5
1. CGPA	1.00				
2. LMS Engagement	0.62	1.00			
3. Attendance	0.58	0.49	1.00		

Variable	1	2	3	4	5
4. Self-Regulation	0.71	0.55	0.47	1.00	
5. Early Prediction Score	0.68	0.74	0.61	0.59	1.00

p < .01

5.5. Student Performance Outcomes

Overall, AI-enhanced predictions reached an AUC of 0.87 and an overall accuracy of 82.4% identifying at-risk students. Participating at-risk students had 14.6 percentage point higher final course pass rates than historical controls who were making the same or no gains as a result of early intervention recommendations based on the framework. The results of the subgroup analysis indicated a minor level of lower predictive equity for students with lower socioeconomic status, which aligns with the more widespread issue of bias in educational AI systems (Gándara et al., 2024).

The overall findings here show that AI-assisted learning analytics tools, based on educational theories and multiple types of data sources, can provide valuable support for predicting students' performance and intervening into it. There were strong relationships between behavioural indicators and outcomes, vouching for the proposed conceptual framework. However, differences between disciplines and demographic groups as well as equity gaps highlight issues in the literature that point to the flaws in using the 'one size fits all' approach and the need for culturally responsive and ethically considered practices (Prinsloo & Slade, 2017; Rets et al., 2023; Holmes, 2023). The results provide valuable insights into the state of AI-LA integration and its prospects for future, as well as actionable insights, and a need for ongoing development and refinement of both the technical and pedagogical aspects of AI-LA integration.

6. Discussion

The results of this study offer insights into the application of AI in learning analytics to meaningfully augment the prediction of student performance and show important nuances in learning practice. AI-powered features can effectively enhance the collection, integration, and interpretation of educational data, transforming static data reports to actionable insights to inform real-time pedagogical interventions (Ouyang et al., 2023; Ahmed et al., 2025). The work revealed that processed indicators including temporal engagement patterns and proxies for self-regulation could be very powerful indicators when used to make predictions, showing that it's not just the quantity of data that's important, but the quality and contextual relevance of the behavioral signals used for prediction.

Among these learning behaviours, LMS interaction consistency in the first half of the semester, forum interaction score and deep approach to forum emerged as the most informative predictors of student success. These behaviors are closely related to the theory and principles of self-regulated learning theory and its use of planning, monitoring, and reflection in an iterative manner (Ifenthaler & Yau, 2020). The strong influence of self-regulation highlights a key research gap: there are many instances of AI-LA implementations which focus heavily on digital traces that are easily measurable, while neglecting metacognitive and motivational aspects that play a crucial role in deep learning (Gašević et al., 2016; Guevara-Reyes et al., 2025). The findings of this study indicate that learner trajectories can be described more fully with hybrid approaches of traced data and validated self-report questionnaires.

This is for teaching practice. AI findings can be used for tailored feedback and personalized interventions, moving away from uniformity and focusing on student-centered teaching that learns from the student. There is evidence to be considered when designing the curriculum, such as what activities most positively relate to success in different disciplines and more intentional scaffolding of self-regulation skills. The use of early identification systems that enable proactive, rather than reactive, interventions, can enhance student support services, thereby avoiding follow-through costs for retention, improving equity outcomes. At the institutional level, it relates to the development of policy with regards to data governance, data resource allocation, and Data Literacy for faculty (Ifenthaler, 2021; OECD, 2026).

Linkages with previous studies yield reinforcement of the importance of integrated frameworks; gaps remain. The results in the present study delve into the same area as Ouyang et al. (2023) and also confirm the effectiveness of AI prediction and the visualization of learning analytics to promote student engagement. Unlike many technically oriented research, this study highlights pedagogical and contextual mediators in an effort to set aside critiques that many learning analytics research is lacking in theory (Gašević et al., 2022). This equity gap in prediciton accuracy between socioeconomic groups was reported elsewhere with respect to algorithmic biases in educative technologies and highlights that fairness-aware modeling is essential (Gándara et al., 2024).

Ethical issues always play an important role in accountable use. While transparency and stakeholder engagement were attempted in the study, the possibility of unintended consequences as increased surveillance anxiety or stigmatization of 'at risk' students signals continuing tensions between predictiveness and autonomy of learners (Prinsloo & Slade, 2017; Holmes, 2023; Rets et al., 2023). The results call for the adoption of by-design methods that not only incorporate care ethics and participative governance but also regard these as essential core elements instead of talking thorns. In conclusion, the dialogue underscores the transformative potential of AI-powered learning analytics, emphasizing the need to combine technological sophistication with pedagogical insights, cultural understanding, and a steadfast dedication to equity and human dignity to achieve success in higher education.

7. Conclusion

The purpose of this research is to develop and test a theoretically supported AI-integrated LA model to predict student performance in HE. The research showed that by combining AI with extensive learning analytics processes and educational theory, it is possible to align and deliver high predictive accuracy along with relevant interventions. The key findings related to the predictive value of combined behaviour and self-regulation indicators, the mediating influence of learner agency and the possibilities of multi-institutional implementation in a variety of contexts.

Overall, the study advances the fields of learning analytics and higher education research by providing an insightful connection between the technical aspects of innovation and pedagogical relevance, where learning analytics research has previously been criticized for lacking in-depth theoretical thinking and limiting the diversity of contexts (Gašević et al., 2016; Ifenthaler, 2021). It helps to contribute conceptual models that emphasize cycles of prediction and support, which are iterative, human centered and provide a basis for the future interdisciplinary scholarship.

In practice, these insights can be used by universities to improve their teaching effectiveness, improve support options for students and for developing policies and practices with data that will ensure excellence and equity. Faculty capacity building and educational ethics mechanisms are encouraged to leverage AI to support, not replace, educational values for the benefit of the student.

The study presents coping loopholes such as dependence on the technological infrastructure of the participating institutions and issues to fully track offline learning experiences. Future work should investigate the longitudinal effects in wider geographical regions and expand the use of multimodal data sources as well as include more advanced fairness measures for AI-LA systems. Parts of the potential impact of these technologies cannot be fully realized without continued critical inquiry, and in the process the humanistic values of higher education will need to be protected.

References

Ahmed, W., et al. (2025). Machine learning-based academic performance prediction with explainability for enhanced decision-making in educational institutions. Scientific Reports, 15, Article 26879. https://doi.org/10.1038/s41598-025-12353-4

Ajayi, O. E. (2026). AI-driven learning analytics in higher education—Insights from a narrative review. Multidisciplinary Digital Publishing Institute Proceedings, 5(1), 16.

Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. IEEE Transactions on Learning Technologies, 10(4), 405–418. https://doi.org/10.1109/TLT.2017.2740172

Boujmiraz, S. (2026). Predicting student performance: A comprehensive review. Computers and Education: Artificial Intelligence. https://doi.org/10.1016/j.caeai.2026.100093

Chukwu, C. O. (2026). AI-powered learning tools on measurement of student engagement across academic disciplines. Educational Point Journal.

Gándara, D., et al. (2024). Algorithmic bias in student success prediction. AERA Open.

Gašević, D., Dawson, S., & Siemens, G. (2016). Let's not forget: Learning analytics are about learning. TechTrends, 59(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Gašević, D., Tsai, Y.-S., & Drachsler, H. (2022). Learning analytics in higher education – Stakeholders, strategy and scale. The Internet and Higher Education, 52, Article 100829.

Guevara-Reyes, R., et al. (2025). Machine learning models for academic performance prediction: Interpretability and application in educational decision-making. Frontiers in Education, 10, Article 1632315. https://doi.org/10.3389/feduc.2025.1632315

Hidayatulloh, W. (2026). Explainable artificial intelligence-based model for student academic performance prediction. Journal of Information System Exploration and Research.

Holmes, W. (2023). Stakeholder perspectives on the ethics of AI in distance-based higher education. International Review of Research in Open and Distributed Learning, 24(4).

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542–570.

Ifenthaler, D. (2021). Putting learning back into learning analytics: Actions for policy makers, practitioners, and researchers. Educational Technology Research and Development, 69, 1–20. https://doi.org/10.1007/s11423-020-09909-8

Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. Educational Technology Research and Development, 68, 1961–1990. https://doi.org/10.1007/s11423-020-09788-5

OECD. (2026). OECD digital education outlook 2026. OECD Publishing.

Ouyang, F., et al. (2023). 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, 20(1), 4. https://doi.org/10.1186/s41239-022-00372-4

Prinsloo, P., & Slade, S. (2017). Big data, higher education and learning analytics: Beyond justice, towards ethics of care. In Big data and learning analytics in higher education (pp. 109–124). Springer.

Raji, B. (2026). AI-enabled predictive analytics in education: Enhancing student success and retention. Journal of Digital Education and Technology.

Rets, I., et al. (2023). Six practical recommendations enabling ethical use of learning analytics. Journal of Learning Analytics, 10(1), 1–20.

Shoaib, M., et al. (2024). AI student success predictor: Enhancing personalized feedback in higher education. Computers in Human Behavior, 158, Article 108292.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. American Behavioral Scientist, 57(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 57(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Wang, Y., et al. (2025). Artificial intelligence in student management systems. PMC Article.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0