



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.935

Volume 15, Issue 10, October 2026

LearnInsight: A Comprehensive Review of Student Performance Analytics Techniques for Learning Gap Detection in Classroom Learning Systems

Ayush Verma, Om Thombre, Chetna Khapre, Alok Dubey, Durga Kashyap, Nikesh Aote

Department of Computer Science & Engineering, Priyadarshini College of Engineering, Nagpur, India

Asst. Professor, Department of Computer Science & Engineering, Priyadarshini College of Engineering, Nagpur, India

ABSTRACT: University courses with hundreds of students have become common, particularly during the early years of university studies. The sheer scale of these courses limits traditional instruction, shifting it towards a one-to-many mode of delivery. This shift reduces student-instructor interaction and the tailored instructor feedback that is crucial for student success. Automated feedback systems allow feedback to be scaled, but they often reduce the instructor's contribution to student learning. This paper investigates how emerging technologies can support, rather than replace, instructors in tailoring their teaching and feedback to identify and correct student knowledge gaps at scale. However, effective classroom learning analytics requires integrating data from multiple digital technologies and interpreting them through domain-specific models of competence development (learning progressions). Such models connect learning processes to long-term competence development, enabling teachers to monitor students' trajectories and adapt instruction to individual needs. Through this task we extract knowledge that describes students' performance in the end-semester examination. It helps in identifying dropouts and students who need special attention earlier, and allows the teacher to provide appropriate advising and counseling.

KEYWORDS: Artificial Intelligence; Large Language Models; Generative Artificial Intelligence; Interfaces; Feedback; Discussion Forums; Learning Analytics; Digital Technologies; Individualized Instruction.

I. INTRODUCTION

The increasing scale of higher education presents a significant pedagogical challenge, straining traditional models of instruction (Ryan et al., 2021). As student numbers in individual courses grow into the hundreds or thousands, particularly in Computer Science (Camp et al., 2017), instructors turn to pedagogies of scale (Ferguson and Sharples, 2014), such as one-to-many lectures and hybrid formats. This shift reduces the capacity for student-instructor interaction and lowers the amount of tailored instructor feedback (Pardo et al., 2019). Both the reduction of teaching tailored to the needs of all students and the diminished possibilities of instructor feedback are problematic. Instructor feedback, unlike that from peers or algorithms, can offer nuanced guidance that corrects deep-seated student knowledge gaps (Banihashem et al., 2024). Importantly, explanations of domain-specific issues provided by instructors show students the ways of thinking foundational to their discipline (Planar and Moya, 2016; Istenič, 2021), which cannot easily be substituted by emergent forms of student support at scale.

The integration of adaptive learning courseware can help close this gap and ensure equitable learning across diverse student groups. Adaptive learning is an approach to personalized learning that moves education away from a "onesize-fits-all" model to one that meets the needs of each learner individually, and it is recognized as having the potential to improve student outcomes by making personalized learning scalable. AI-driven adaptive feedback systems have emerged as a promising solution to this challenge. These systems use learning analytics and machine learning algorithms to analyze student interactions, detect misunderstandings, and provide real-time personalized feedback. The effectiveness of AI-driven feedback in closing conceptual learning gaps remains an area requiring empirical

II. RELATED WORK

A. Identifying Student Knowledge Gaps

Delivering personalized and effective instruction to a large number of students remains a significant hurdle in educational settings. Existing approaches have tried to address this by scaling instruction via video and quizzes (Cummins et al., 2015; Mirriahietal.,2021) and by improving engagement through new pedagogies of scale (Davis etal., 2018) [1].

Early approaches to identifying students' knowledge gaps, such as quizzes with closed- or open-ended questions, often require human labeling and lack the responsiveness needed to offer feedback when students need it most. Stephens-Martinez et al. suggested a semi-automated method requiring minimal human labeling, which predicted knowledge gaps through statistical co-occurrence patterns and relied on recall and precision values [2]. There is a critical need for technologies that support both the quality and the impact of instructor-led instruction, so that students receive the guidance essential for effective learning and correction of knowledge gaps.

B. LLMs for Generating Learning Opportunities

A growing body of work demonstrates how LLMs can enhance the automatic generation of learning opportunities. Denny et al. (2023) examined whether LLMs could generate educational resources comparable to those produced by computer science students and found that LLM-generated resources were less variable in length and more consistent with existing examples.

Chan et al. (2025) investigated how different prompting methods influence quality and accuracy when leveraging GenAI to produce learning opportunities. They found that Chain-of-Prompting with code prompting yielded superior results compared to standard prompting, while noting high variance across STEM subjects. In contrast to fully automated generation, Choi et al. (2024) suggested a method that enables instructors to use LLMs to transform lecture transcripts into educational dialogues. This approach requires instructors to guide the AI through multi-step processes such as generating rubrics and questions. While LLM capabilities are promising, challenges such as hallucination, where the AI produces inaccurate content, remain

TABLE I. SUMMARY OF RELATED STUDIES

Study	Year	Approach	Key Finding
Davis et al. [1]	2018	Review of online learning strategies at scale	Personalized instruction for large cohorts remains a hurdle
Stephens-Martinez et al. [2]	2016	Semi-automated, cooccurrence based	Predicts knowledge gaps with minimal human labeling
Denny et al.	2023	LLM-generated learning resources	More consistent and less variable in length
Chan et al.	2025	Prompting methods for GenAI	Chain-of-Prompting with code best; high variance across STEM
Choi et al.	2024	LLM-based lecture transcript to dialogue	Needs instructor guidance; hallucination remains a risk
Aya et al.	—	DNN, RF, SVM, KNN	DNN best, 89% accuracy
Bravo-Agapito et al. [9]	2021	Regression, factor and cluster analysis on Moodle data	Forum visits, quizzes, assignments and age most influential
Pandey and Pal [7]	2011	Bayes classification	Predicts whether new students will perform well
Neumann et al. [11]	2024	Overview of GenAI in education	Mostly supports teaching, rarely analyzes learning
Jin et al. [12]	2024	Learning progressions	Domain models guide longterm competence development

C. Predicting Student Performance

Aya et al. studied the prediction of students' successful completion of courses such as "programming" and "data structures", which students fail easily. Using a publicly available student achievement dataset from a four-year university, they compared deep neural networks (DNNs), random forests, support vector machines (SVMs), and Knearest neighbors (KNN). The DNN gave the best prediction, with an accuracy of 89%, further demonstrating the effectiveness of DNNs in predicting academic performance.

Bravo-Agapito et al. noted that the online teaching model has developed rapidly and that large amounts of students' online learning information are now recorded in detail. They constructed a series of models to predict early student achievement using Moodle interaction data and performance characteristics from 802 undergraduate students at CO University, including exploratory factor analysis, multiple linear regression, cluster analysis, and correlation analysis. The four factors with the greatest impact on achievement were the number of visits to forums and glossaries, questionnaires, assignments, and student age [9].

Pandey and Pal studied student performance by selecting 600 students from different colleges of Dr. R. M. L. Awadh University, Faizabad, India. Using Bayes classification on category, language, and background qualification, they identified whether newcomer students would perform well or not [7].

D. Generative AI and Learning Progressions

Generative AI techniques have only very recently been recognized as a means to analyze and support student learning and are still subject to research (for an overview see Neumann et al., 2024) [11]. Because of its ability to generate new content, much of the research in education on generative AI has revolved around supporting teaching rather than analyzing student learning.

Classroom instruction in the 21st century is intended to support students in developing competence. This requires carefully orchestrated learning opportunities over longer periods of time, and therefore domain-specific models that can guide this orchestration across one grade, multiple grades, and even grade bands (Jin et al., 2024) [12]. Such models have only recently become available at scale.

III. CONCLUSION

This work addresses the problem of limited instructor feedback in large classes, which stems from their sheer scale. To tackle this, we propose an approach that leverages discussion forums. This dual-purpose solution aims to identify students' knowledge gaps to meet their authentic needs and simultaneously provide personalized support through tailored learning opportunities. Crucially, the approach does not seek to replace instructors' judgment but to complement it by providing actionable insights where decision-making is difficult due to scale and incomplete evidence. Our approach is demonstrated through an instructor-facing educational tool that combines LLMs and RAG. This tool identifies common classlevel knowledge gaps and generates targeted learning activities.

This study will help students and teachers improve student learning and performance. It will also help identify students who need special attention, reduce failure rates, and enable instructors to take appropriate actions in preparation for the next semester's examinations.

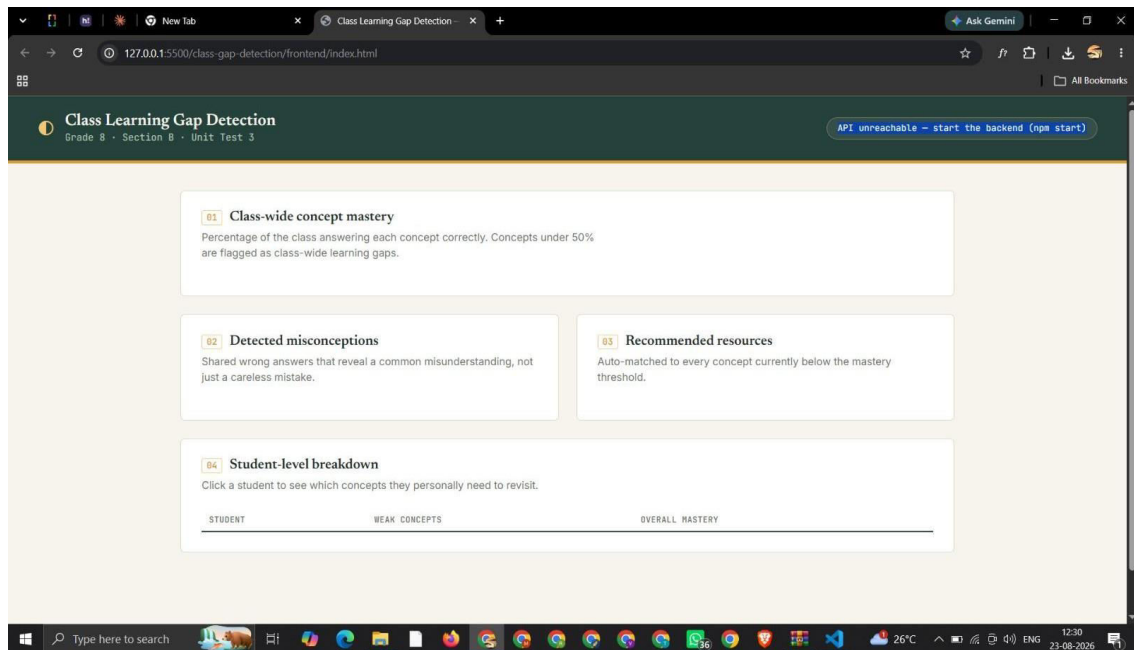
IV. RESEARCH GAP

- Existing systems focus mainly on marks/grades and do not sufficiently analyze the reasons behind students' learning gaps.
- Existing analytics systems often report performance rather than predict or identify emerging learning gaps.
- Many existing approaches do not provide personalized or topic-wise insights about which concepts a student has not mastered.
- A need exists for a simple, automated, teacher-friendly system that can identify students or topics requiring additional attention.

REFERENCES

1. D. Davis, G. Chen, C. Hauff, and G.-J. Houben, "Activating learning at scale: A review of innovations in online learning strategies," *Computers & Education*, vol. 125, pp. 327–344, 2018, doi: 10.1016/j.compedu.2018.05.019.
2. K. Stephens-Martinez, A. Ju, C. Schoen, J. DeNero, and A. Fox, "Identifying student misunderstandings using constructed responses," in *Proc. Third ACM Conf. Learning at Scale*, 2016, pp. 153–156.
3. T. Song, H. Zhang, and Y. Xiao, "A high-quality generation approach for educational programming projects using LLM," *IEEE Trans. Learning Technologies*, 2024.
4. P. A. Mane and D. S. Jagtap, "AI-enhanced educational platform for personalized learning paths, automated grading, and real-time feedback," *Int. J. Sci. Res. Eng. Manag.*, vol. 8, pp. 1–7, 2024.
5. A. S. Alsolami, "The effectiveness of using artificial intelligence in improving academic skills of school-aged students with mild intellectual disabilities in Saudi Arabia," *Res. Dev. Disabil.*, vol. 156, art. 104884, 2025.
6. H. Xie, H.-C. Chu, G.-J. Hwang, and C.-C. Wang, "Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017," *Comput. Educ.*, vol. 140, art. 103599, 2019.
7. U. K. Pandey and S. Pal, "A data mining view on class room teaching language," *Int. J. Computer Science Issues (IJCSI)*, vol. 8, no. 2, pp. 277–282, 2011, ISSN: 1694-0814.
8. H. Mannila, *Data Mining: Machine Learning, Statistics, and Databases*. IEEE, 1996.
9. J. Bravo-Agapito, S. J. Romero, and S. Pamplona, "Early prediction of undergraduate student's academic performance in completely online learning: A five-year study," *Computers in Human Behavior*, vol. 115, art. 106595, 2021.
10. S. Hussain et al., "Regression analysis of student academic performance using deep learning," *Education and Information Technologies*, vol. 26, no. 1, pp. 783–798, 2021.
11. K. Neumann, J. Kuhn, and H. Drachler, "Generative Künstliche Intelligenz in Unterricht und Unterrichtsforschung – Chancen und Herausforderungen," *Unterrichtswissenschaft*, vol. 52, no. 2, pp. 227–237, 2024.
12. H. Jin, D. Yan, and J. S. Krajcik, Eds., *Handbook of Research on Science Learning Progressions*. New York: Routledge, 2024.
13. Z. Zhang, Z. Li, H. Liu, T. Cao, and S. Liu, "Data-driven online learning engagement detection via facial expression and mouse behavior recognition technology," *J. Educational Computing Research*, vol. 58, no. 1, pp. 63–86, 2020.
14. S. Chan, N. Lo, and A. Wong, "Leveraging generative AI for enhancing university-level English writing: Comparative insights on automated feedback and student engagement," *Cogent Educ.*, vol. 12, art. 2440182, 2024.
15. Z. Zhang, Z. Li, H. Liu, T. Cao, and S. Liu, "Data-driven online learning engagement detection via facial expression and mouse behavior recognition technology," *J. Educational Computing Research*, vol. 58, no. 1, pp. 63–86, 2020.
16. R. Bertolini, S. J. Finch, and R. H. Nehm, "Testing the impact of novel assessment sources and machine learning methods on predictive outcome modeling in undergraduate biology," *J. Science Education and Technology*, vol. 30, no. 2, pp. 193–209, 2021.
17. M. Millea, R. Wills, A. Elder, and D. Molina, "What matters in college student success? Determinants of college retention and graduation rates," *Education*, vol. 138, pp. 309–322, 2018.
18. L. Konnova, L. Lipagina, G. Postovalova, A. Rylov, and I. Stepanyan, "Designing adaptive online mathematics course based on individualization learning," *Educ. Sci.*, vol. 9, art. 182, 2019.
19. I. Molenaar and C. A. Knoop-van Campen, "How teachers make dashboard information actionable," *IEEE Trans. Learning Technologies*, vol. 12, pp. 347–355, 2018.
20. K. Taneja, P. Maiti, S. Kakar, P. Guruprasad, S. Rao, and A. K. Goel, "Jill Watson: A virtual teaching assistant powered by ChatGPT," in *Int. Conf. Artificial Intelligence in Education*, Springer, 2024, pp. 324–337.
21. D. S. Fleur, W. van den Bos, and B. Bredeweg, "Social comparison in learning analytics dashboards supporting motivation and academic achievement," *Computers and Education Open*, vol. 4, art. 100130, 2023.

OUTPUT





INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details