Tavhare Pranjal Pravin, 2026, 14:5
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Ai In Education: A Review Of Multimodal Student Engagement Detection, Learning Analytics And Adaptive Intervention Systems

Tavhare Pranjal Pravin[1] Prof. S. S. Bhor[2] Prof. M. C. Kolse[3] Prof. S. K. Said[4]
Department of Artificial Intelligence and Data Science
Jaihind College of Engineering
Pune, Maharashtra, India

Abstract—Student engagement is closely linked to achievement, retention and the quality of learning, yet teachers in large physical classrooms and online courses cannot easily monitor every learner in real time. This paper reviews AI-driven student engagement systems for classroom and online settings. It begins with the educational view of engagement as behavioural, emotional and cognitive, and then examines automatic engagement recognition from facial expressions and head pose, datasets of learners recorded in the wild, the dynamics of affective states such as confusion and boredom during learning, and learning analytics from activity logs, attendance and quiz performance, including studies of disengagement patterns in massive open online courses and knowledge tracing. Machine learning and deep learning methods such as CNNs, random forests and SVMs for classifying engagement as high, moderate or low are compared using accuracy, precision and recall. An adaptive recommendation module that suggests content and interventions and a teacher dashboard are discussed. A five-layer architecture comprising multimodal data, feature extraction, engagement classification, adaptive intervention and teacher analytics layers is proposed, together with research gaps, objectives, a comparison of data modalities and challenges such as privacy, consent, bias and scalability. The review concludes that ethical, teacher-centred AI can support more inclusive and student-centred education.
Index Terms—Student engagement, AI in education, learning analytics, facial expression recognition, engagement detection, personalized learning, teacher dashboard

I. INTRODUCTION

In a classroom of sixty students, a teacher may notice that a few are distracted, but cannot track each learner's attention throughout a lecture. In online courses, the problem is worse: a student may keep a video playing while doing something else. Low engagement often precedes poor results and dropout.

Fredricks *et al.* [1] described engagement as having behavioural (participation), emotional (interest and feelings) and cognitive (investment in learning) dimensions. AI can estimate these dimensions from faces, interactions and performance, giving teachers early signals. Siemens and Baker [2] described learning analytics and educational data mining as complementary fields using data to improve learning.

This paper reviews AI methods for engagement detection and intervention and proposes an ethical, teacher-centred architecture.

II. THEORETICAL BACKGROUND

1. Dimensions And Signals Of Engagement

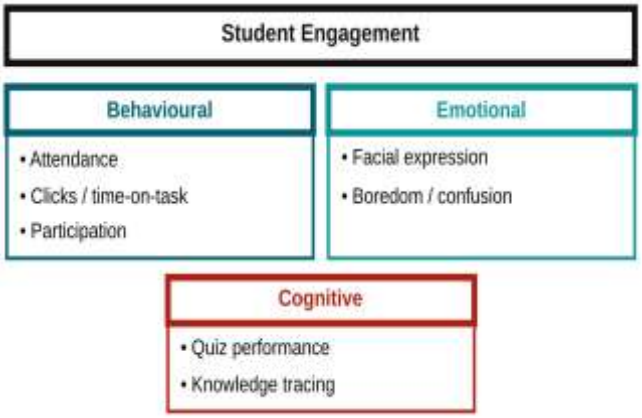


Fig. 1 maps the three dimensions of engagement to observable signals and AI methods.

Fig. 1. Dimensions of student engagement and corresponding signals.

© 2026 Tavhare Pranjal Pravin, This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

2. Engagement Levels

Systems usually classify engagement into ordered levels, such as highly engaged, moderately engaged and disengaged, over short time windows.

III. SURVEY OF EXISTING APPROACHES

Visual engagement recognition. Whitehill *et al.* [3] showed that machine learning on facial expressions can recognize student engagement with accuracy comparable to human observers. Gupta *et al.* [4] released DAiSEE, a dataset of students watching online content in uncontrolled settings, labelled for engagement, boredom, confusion and frustration. Facial expression recognition methods are surveyed by Li and Deng [5].

Affective dynamics. D'Mello and Graesser [6] studied how confusion, frustration and boredom evolve during complex learning, showing that confusion can be productive if resolved but leads to boredom if not.

Learning analytics. Kizilcec *et al.* [7] clustered MOOC learners into completing, auditing, disengaging and sampling groups based on activity patterns. Kuzilek *et al.* [8] released the Open University learning analytics dataset, and deep knowledge tracing [9] models student knowledge from answer sequences.

Classifiers. Random forests [10] and SVMs [11] are common for log-based features, while CNNs process faces. Table I summarizes these works.

TABLE I Summary of Literature Survey

Sr. No.	Ref.	Year	Methodology	Findings	Research Gap
1	[1]	2004	Engagement concept review	Three dimensions	Not computational
2	[6]	2012	Affective dynamics	Confusion–boredom patterns	Lab tasks
3	[7]	2013	MOOC disengagement	Learner subpopulations	Log data only
4	[3]	2014	Facial engagement recognition	Near-human accuracy	Small sample
5	[9]	2015	Deep knowledge tracing	Models student knowledge	Interpretability
6	[4]	2016	DAiSEE dataset	In-the-wild labels	Label imbalance
7	[8]	2017	OULAD dataset	Rich learning logs	Single university

IV. COMPARATIVE STUDY

Table II compares data modalities. Logs are scalable and privacy-friendly, faces add real-time emotional cues, and quizzes reflect cognitive engagement but less frequently.

Modality	Dimension	Real-time	Privacy Risk
Facial video	Emotional / behavioural	Yes	High
LMS activity logs	Behavioural	Near	Low
Quiz performance	Cognitive	No	Low
Attendance	Behavioural	No	Low
Multimodal fusion	All three	Yes	Medium

TABLE II Comparison of Engagement Data Modalities

V. RESEARCH GAPS

1. 1) Facial cues do not always reflect cognitive engagement.

2. 2) Datasets are small and lack diversity, including Indian classrooms.

3. 3) Continuous camera monitoring raises privacy and consent concerns.

4. 4) Detection is rarely connected to effective interventions.

5. 5) Models may be biased across gender, skin tone and culture.

VI. AIMS AND OBJECTIVES

- **O1:** To review the concept and dimensions of student engagement.
- **O2:** To examine visual, behavioural and performance-based engagement detection.
- **O3:** To compare machine learning and deep learning classifiers.
- **O4:** To propose a five-layer engagement and intervention architecture.
- **O5:** To discuss privacy, ethics and scalability.

VII. PROPOSED ARCHITECTURE

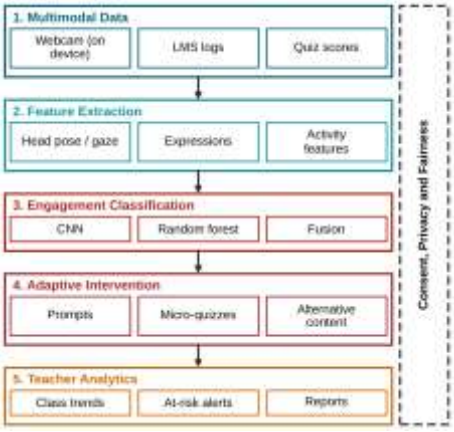


Fig. 2 shows the proposed architecture.

Fig. 2. Proposed five-layer AI-driven student engagement architecture.

1. Multimodal Data Layer

With consent, the system collects webcam frames (processed on device), attendance, LMS activity logs, quiz scores and forum interactions.

B. Feature Extraction Layer

Head pose, gaze direction and expression features are extracted from video; activity frequency, time-on-task and score trends from logs.

2. Engagement Classification Layer

A CNN for visual features and a random forest for log features are fused to classify engagement level per time window.

3. Adaptive Intervention Layer

Disengaged students receive prompts, short quizzes, alternative explanations or peer activities; repeated disengagement is flagged to teachers.

4. Teacher Analytics Layer

A dashboard shows class-level engagement trends and students needing support, without exposing raw video.

VIII. PROCESS FLOW AND EVALUATION

Fig. 3 describes the workflow. Visual and log-based probabilities are fused as

$$p = \alpha \cdot p_visual + (1 - \alpha) \cdot p_logs \quad (1)$$

and the engagement level is the class with the highest fused probability. Performance is measured by accuracy, precision and recall,

$$Precision = TP / (TP + FP), \quad Recall = TP / (TP + FN) \quad (2)$$

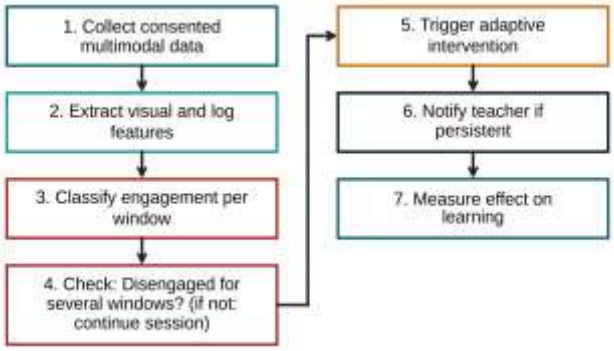


Fig. 3. Engagement detection and intervention workflow.

IX. LIMITATIONS AND FUTURE SCOPE

Table III summarizes challenges. Future directions include on-device processing for privacy, LLM-based tutoring that responds to engagement, support for Indian languages and evaluation of long-term learning gains.

Challenge	Mitigation
Privacy of video	On-device processing; no storage
Consent	Opt-in; transparent policy
Bias	Diverse datasets; fairness tests
Misinterpretation	Teacher judgement; context
Scalability	Lightweight models; cloud LMS

TABLE III Challenges and Mitigation

X. CONCLUSION

AI-based engagement systems can help teachers identify struggling learners early and personalize support. The reviewed literature shows that faces, interaction logs and performance data each provide useful signals, and that combining them improves detection. Privacy-preserving, transparent and teacher-centred designs are essential so that these systems support learners rather than surveil them.

1. Acknowledgment

The authors express their sincere gratitude to the Department of Artificial Intelligence and Data Science, Jaihind College of Engineering, Pune, for providing the facilities, guidance and encouragement needed to carry out this work. The authors also thank the faculty members and colleagues whose valuable suggestions helped improve the quality of this review.

REFERENCES

1. J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, "School engagement: Potential of the concept, state of the evidence," Rev. Educ. Res., vol. 74, no. 1, pp. 59–109, 2004.
2. G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in Proc. 2nd Int. Conf. Learn. Analytics Knowl. (LAK), 2012, pp. 252–254.
3. J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan, "The faces of engagement: Automatic recognition of student engagement from facial expressions," IEEE Trans. Affect. Comput., vol. 5, no. 1, pp. 86–98, 2014.
4. A. Gupta, A. D'Cunha, K. Awasthi, and V. Balasubramanian, "DAiSEE: Towards user engagement recognition in the wild," arXiv:1609.01885, 2016.
5. S. Li and W. Deng, "Deep facial expression recognition: A survey," IEEE Trans. Affect. Comput., vol. 13, no. 3, pp. 1195–1215, 2022.
6. S. D'Mello and A. Graesser, "Dynamics of affective states during complex learning," Learn. Instr., vol. 22, no. 2, pp. 145–157, 2012.
7. R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses," in Proc. 3rd Int. Conf. Learning Analytics Knowl. (LAK), 2013, pp. 170–179.
8. J. Kuzilek, M. Hlosta, and Z. Zdrahal, "Open University learning analytics dataset," Sci. Data, vol. 4, Art. no. 170171, 2017.
9. C. Piech et al., "Deep knowledge tracing," in Proc. NeurIPS, vol. 28, 2015.
10. L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.
11. C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, 1995.