American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)

E-ISSN: XXXX-XXXX ● Website: www.ajamri.com ● Email: editor@ajamri.com

Intelligent Learning Analytics: Using Multimodal Data to Understand Student Engagement and Academic Progress

Cindy Nakatsu

Professor Purdue University

Abstract

The increasing adoption of digital learning environments has generated large volumes of educational data that can be used to understand student behaviour, engagement and academic progress. Traditional learning analytics primarily relies on structured indicators such as grades, attendance and assessment scores, whereas intelligent learning analytics increasingly incorporates multimodal data derived from learning management systems, classroom interactions, assessment responses, language, video, audio, eye-tracking, clickstream activity and other digital traces. This paper examines how multimodal data and artificial intelligence can support a more comprehensive understanding of student engagement and academic progress. A conceptual framework is proposed that integrates data acquisition, preprocessing, multimodal feature extraction, artificial intelligence-based analytics, learner profiling, early-warning mechanisms and personalised interventions. The paper discusses how behavioural, cognitive, emotional and social dimensions of engagement can be analysed through complementary data sources. It further examines challenges associated with data quality, privacy, algorithmic bias, interpretability, consent and the risk of excessive student surveillance. Particular attention is given to the distinction between observable digital behaviour and genuine learning, emphasising that increased platform activity does not necessarily indicate deeper engagement or improved academic achievement. The paper proposes a human-centred intelligent learning analytics framework in which AI-generated insights support rather than replace teacher judgement. The study concludes that multimodal learning analytics can strengthen early identification of learning difficulties, personalised feedback and evidence-based instructional design when implemented with appropriate pedagogical validation, ethical safeguards and transparent governance.

Keywords: Learning Analytics, Multimodal Data, Student Engagement, Academic Progress, Artificial Intelligence, Educational Data Mining, Learning Management Systems, Personalised Learning, Predictive Analytics, Student Success.

1. Introduction

Digital transformation has fundamentally changed the way educational activities are delivered, monitored and evaluated.

Modern learning environments generate extensive data through:

- Learning management systems.
- Online assessments.
- Digital classrooms.
- Discussion forums.
- Video-learning platforms.
- Educational applications.
- Student information systems.

These data provide opportunities to understand student learning beyond traditional examination results.

Conventional educational evaluation often focuses on:

Grades + Attendance + Assessment Scores

Intelligent learning analytics can extend this model towards:

Behaviour + Interaction + Assessment + Language + Context + Learning Outcomes

The objective is not simply to collect more data but to extract meaningful information that can improve learning and teaching.

2. Learning Analytics

Learning analytics refers broadly to the collection, measurement, analysis and interpretation of data about learners and learning environments.

Its objectives can include:

- Understanding learning behaviour.
- Identifying academic difficulties.
- Predicting student outcomes.
- Supporting personalised learning.
- Improving instructional design.

3. Intelligent Learning Analytics

Intelligent learning analytics combines learning analytics with artificial intelligence and advanced data-processing methods.

Potential technologies include:

- Machine learning.
- Natural language processing.
- Deep learning.
- Pattern recognition.
- Predictive modelling.

- Recommender systems.

These technologies can identify patterns that may not be immediately visible through conventional reporting.

4. Multimodal Learning Data

Multimodal data refers to information collected through different forms or channels.
Potential sources include:

Data Modality	Examples
Behavioural	Clicks, navigation, time-on-task
Academic	Grades, quiz scores, assignments
Textual	Discussion posts, written responses
Audio	Classroom or presentation speech
Visual	Video interactions, facial cues
Physiological	Eye tracking or other signals
Social	Peer interaction and collaboration
Contextual	Device, timing, learning environment

The combination of these modalities can provide a richer learner profile.

5. Student Engagement

Student engagement is multidimensional.
It commonly includes:

Behavioural Engagement

Participation in learning activities.

Cognitive Engagement

Mental effort and investment in learning.

Emotional Engagement

Interest, motivation and attitudes.

Social Engagement

Interaction with teachers and peers.
A major challenge is that no single data source can accurately represent all these dimensions.

6. Behavioural Engagement

Digital platforms can capture observable behaviours such as:

- Login frequency.
- Resource access.
- Assignment submission.
- Video viewing.
- Discussion participation.
- Time spent on activities.

These indicators can provide useful signals of participation.
However, activity should not automatically be interpreted as learning.

7. Cognitive Engagement

Cognitive engagement is more difficult to measure because it involves internal mental processes.
Potential indicators include:

- Quality of written responses.
- Problem-solving strategies.
- Revision behaviour.
- Question complexity.
- Conceptual understanding.

Natural language processing can potentially analyse textual responses to identify patterns associated with deeper reasoning.

8. Emotional Engagement

Emotional engagement may involve:

- Interest.
- Frustration.
- Confidence.
- Anxiety.
- Motivation.

Some research systems attempt to infer affective states from:

- Language.
- Facial expressions.
- Voice.
- Interaction patterns.

However, emotional inference is highly context-dependent and should be treated cautiously.

9. Social Engagement

Collaborative learning generates social data.
Examples include:

- Peer responses.

- Group discussions.
- Collaborative documents.
- Forum interactions.
- Peer assessment.

Network analysis can help identify participation patterns within learning communities.

10. Academic Progress

Academic progress can be assessed through:

- Assessment scores.
- Assignment quality.
- Learning trajectories.
- Competency achievement.
- Course completion.
- Skill development.

Intelligent analytics can examine progress longitudinally rather than relying only on individual assessments.

11. Learning Trajectories

A student's learning trajectory represents changes in performance and engagement over time.
For example:

Initial Understanding
↓
Practice
↓
Assessment
↓
Feedback
↓
Improvement

Analytics can identify where students deviate from expected trajectories.

12. Early-Warning Systems

AI can potentially identify students who may require additional support.
Possible indicators include:

- Declining assessment performance.
- Reduced participation.
- Repeated failed attempts.
- Increasing inactivity.
- Sudden changes in learning behaviour.

An early-warning system should trigger support, rather than automatically labelling students as unsuccessful.

13. Predictive Learning Analytics

Predictive models may estimate outcomes such as:

- Probability of course completion.
- Risk of academic failure.
- Expected assessment performance.
- Likelihood of disengagement.

Potential models include:

- Logistic regression.
- Decision trees.
- Random forests.
- Gradient boosting.
- Neural networks.

14. Multimodal Data Fusion

The central challenge is integrating different data types.
A conceptual architecture is:

Behavioural Data
•
Academic Data
•
Text Data
•
Interaction Data
↓
Multimodal Data Fusion
↓
AI Analytics
↓
Learner Profile
↓
Personalised Intervention

15. Data Preprocessing

Raw educational data often contain:

- Missing values.
- Duplicate records.
- Inconsistent timestamps.

- Noise.
- Different measurement scales.

Preprocessing is therefore necessary before multimodal analysis.

16. Feature Engineering

Relevant features can be extracted from different modalities.
Examples include:

Behavioural Features

- Session frequency.
- Session duration.
- Resource completion.

Academic Features

- Score trends.
- Error frequency.
- Assessment attempts.

Text Features

- Vocabulary complexity.
- Topic relevance.
- Conceptual indicators.

17. Natural Language Processing

NLP can analyse student-generated text.
Potential applications include:

- Automated feedback.
- Essay analysis.
- Discussion analysis.
- Question classification.
- Concept identification.

NLP should support teachers rather than become the sole basis for high-stakes academic decisions.

18. Computer Vision

Video-based systems can potentially analyse classroom or online-learning interactions.
Possible applications include:

- Attention estimation.
- Gesture analysis.
- Participation patterns.

However, visual analytics raises substantial privacy and ethical concerns.

19. Audio Analytics

Speech data can potentially provide information about:

- Participation.
- Question frequency.
- Discussion patterns.
- Presentation performance.

Audio analytics must account for language, accent and communication differences.

20. Eye-Tracking Data

Eye-tracking can provide information about visual attention.

Potential measures include:

- Fixation duration.
- Saccades.
- Gaze patterns.

Such data may help researchers understand interaction with educational content.

However:

Visual attention ≠ Learning

A learner may look at content without understanding it.

21. Multimodal Engagement Index

A conceptual engagement score can combine multiple indicators.

For example:

Engagement = Behaviour + Cognitive Indicators + Social Interaction + Context

The weights should be empirically validated rather than arbitrarily assigned.

22. Personalised Learning

Intelligent analytics can support personalised learning by identifying:

- Knowledge gaps.
- Preferred learning resources.
- Difficulty patterns.
- Learning pace.

The system can then recommend appropriate:

- Exercises.
- Videos.
- Readings.
- Assessments.

23. Adaptive Learning

Adaptive systems can modify learning pathways according to student performance.
For example:
Low Mastery → Additional Explanation
Moderate Mastery → Guided Practice
High Mastery → Advanced Task
This can reduce unnecessary repetition while providing additional support where needed.

24. Intelligent Feedback

AI can provide rapid feedback on:

- Quizzes.
- Written responses.
- Coding exercises.
- Problem-solving tasks.

Effective feedback should be:

- Specific.
- Constructive.
- Actionable.
- Timely.

25. Teacher Decision Support

Learning analytics can help teachers identify:

- Students requiring support.
- Difficult concepts.
- Participation gaps.
- Assessment weaknesses.

The teacher remains responsible for interpreting the broader context.

26. Teacher-in-the-Loop Architecture

A human-centred system can follow:

Data
↓
AI Analysis
↓
Teacher Dashboard
↓
Teacher Interpretation
↓
Student Intervention
↓
Outcome Evaluation

This reduces the risk of fully automated educational decision-making.

27. Student Dashboards

Learner-facing dashboards can provide:

- Progress indicators.
- Skill mastery.
- Learning goals.
- Activity summaries.
- Feedback.

Students can use these insights for self-regulated learning.

28. Self-Regulated Learning

Learning analytics can support:

1. Goal setting.
2. Monitoring.
3. Reflection.
4. Strategy adjustment.

This can encourage learners to become more active participants in their own learning processes.

29. Recommendation Systems

Educational recommendation engines can suggest learning resources based on:

- Previous performance.
- Learning history.
- Competency gaps.
- Current learning objectives.

However, recommendation systems should avoid creating overly narrow learning pathways.

30. Academic Risk Prediction

Risk models may identify patterns associated with academic difficulty.

Possible outputs include:

- Low-risk.
- Moderate-risk.
- High-risk.

These categories should be interpreted as support signals, not fixed predictions of student ability.

31. Intervention Strategies

When analytics identify potential difficulties, interventions may include:

- Additional tutoring.
- Targeted learning resources.
- Teacher feedback.
- Peer support.
- Academic advising.

The effectiveness of the intervention should subsequently be evaluated.

32. Longitudinal Learning Analytics

Longitudinal analysis examines learner data over extended periods.

It can identify:

- Persistent learning difficulties.
- Improvement trends.
- Engagement cycles.
- Long-term competency development.

This is often more informative than analysing isolated events.

33. Learning Analytics for Online Education

Online education produces extensive digital traces.

These include:

- Video interactions.
- Discussion participation.
- Quiz attempts.
- Download behaviour.
- Navigation sequences.

These data can support detailed analysis of online learning behaviour.

34. Learning Analytics in Blended Learning

Blended learning combines physical and digital education.

Analytics can integrate:

Classroom Data + Digital Data

This creates a more complete representation of learning activity.

35. Learning Analytics in Higher Education

Universities can use analytics for:

- Student retention.
- Course improvement.
- Academic advising.
- Curriculum evaluation.

- Learning-support services.

Institutional implementation should maintain strong governance over student data.

36. Data Privacy

Multimodal analytics can involve sensitive information.

Potential data include:

- Academic performance.
- Behavioural traces.
- Communication.
- Video.
- Audio.
- Physiological signals.

Institutions must establish clear rules for collection, storage and use.

37. Informed Consent

Students should understand:

- What data are collected.
- Why they are collected.
- How they are analysed.
- Who can access them.
- How long they are retained.

Consent processes should be transparent and understandable.

38. Data Minimisation

Institutions should collect only data that are genuinely necessary for educational objectives.

More data do not necessarily produce better analytics.

The principle should be:

Relevant Data > Maximum Data

39. Algorithmic Bias

AI systems may produce biased predictions because of:

- Unrepresentative datasets.
- Historical inequalities.
- Measurement errors.
- Model design.

Bias assessment should therefore be integrated into system development.

40. Equity and Inclusion

Learning analytics should account for differences in:

- Socioeconomic background.

- Disability.
- Language.
- Access to technology.
- Learning environment.

A student with limited connectivity may appear less engaged even when highly motivated.

41. Explainability

Students and educators should be able to understand important analytical outputs.
For example:
"Why was this learner identified as requiring additional support?"
The system should provide understandable contributing factors.

42. False Positives

An AI system may identify a student as academically at risk even when the student is performing adequately.
False positives can result in:

- Unnecessary interventions.
- Student anxiety.
- Stigmatisation.

Human review is therefore essential.

43. False Negatives

The opposite problem occurs when a struggling learner is not identified.
False negatives can be particularly harmful because the learner may not receive timely support.

44. Surveillance Concerns

Continuous monitoring can create a perception of surveillance.
Students may alter their behaviour because they know their:

- Clicks.
- Attention.
- Communication.
- Activity.

are being monitored.
Educational analytics should therefore prioritise learning support over behavioural control.

45. Academic Freedom

Analytics should not unnecessarily constrain students' learning choices.
Systems should support exploration and creativity rather than forcing learners into rigid behavioural patterns.

American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)
E-ISSN: XXXX-XXXX ● Website: www.ajamri.com ● Email: editor@ajamri.com

46. Model Validation

Learning analytics models should be evaluated using:

- Predictive accuracy.
- Precision.
- Recall.
- Calibration.
- Fairness.
- Generalisability.

Performance should be tested across different student populations.

47. Model Drift

Student behaviour and educational environments change.

A model developed for one:

- Course.
- Institution.
- Student population.

may perform differently elsewhere.

Continuous validation is therefore required.

48. Proposed Intelligent Learning Analytics Framework

The proposed framework consists of eight stages:

Stage 1 — Multimodal Data Collection

Collect relevant academic, behavioural and interaction data.

Stage 2 — Data Governance

Apply privacy, consent, security and access controls.

Stage 3 — Data Preprocessing

Clean and standardise multimodal datasets.

Stage 4 — Feature Extraction

Identify meaningful indicators from each modality.

Stage 5 — Multimodal AI Analysis

Combine data streams to identify engagement and progress patterns.

Stage 6 — Learner Profiling

Develop dynamic and interpretable learning profiles.

Stage 7 — Human-Centred Intervention
Teachers and advisors use insights to provide appropriate support.

Stage 8 — Outcome Evaluation
Measure whether interventions improve learning outcomes.

49. Conceptual Architecture

Student
↓
Digital Learning Environment
↓
Multimodal Data

- Behaviour
- Assessment
- Text
- Audio
- Video
- Social Interaction

↓
Data Governance & Preprocessing
↓
Multimodal AI Engine
↓
Engagement + Progress Analysis
↓
Teacher Dashboard + Student Dashboard
↓
Personalised Intervention
↓
Learning Outcome
↓
Continuous Feedback

50. Research Questions

1. How can multimodal data improve the measurement of student engagement?
2. Which combinations of data modalities provide the most reliable indicators of academic progress?
3. How accurately can AI predict student learning difficulties?
4. Can multimodal analytics improve personalised learning recommendations?

5. How can learning analytics distinguish participation from genuine learning engagement?
6. What ethical risks emerge from continuous multimodal student monitoring?
7. How can explainable AI improve trust in learning analytics?
8. What role should teachers play in AI-supported educational decision-making?
9. How can institutions minimise algorithmic bias in student-risk prediction?
10. What governance mechanisms are necessary for responsible multimodal learning analytics?

51. Future Research Directions

Future research should investigate:

- Multimodal foundation models for education.
- AI-based learning companions.
- Explainable student-risk prediction.
- Privacy-preserving learning analytics.
- Federated educational analytics.
- Emotion-aware learning systems.
- Digital twins of learning environments.
- Adaptive assessment.
- AI-assisted academic advising.
- Cross-cultural learning analytics.
- Longitudinal learner modelling.
- Human–AI collaboration in education.

52. Discussion

Intelligent learning analytics provides an opportunity to move beyond simple educational dashboards.

Traditional systems may tell educators:

“The student scored 55%.”

More sophisticated systems can potentially identify:

“The student has demonstrated declining performance, reduced participation and persistent difficulty with a specific competency.”

This distinction is important because useful educational analytics should provide actionable insight, not simply more data.

Multimodal analysis can strengthen this process by combining different perspectives on learning. Nevertheless, more modalities do not automatically produce more accurate conclusions. Each modality introduces its own measurement limitations, biases and ethical considerations.

For example, reduced online activity may indicate disengagement, but it may also result from:

- Internet access problems.

- Employment commitments.
- Family responsibilities.
- Alternative learning strategies.

Therefore, contextual interpretation remains essential.

53. From Predictive Analytics to Prescriptive Learning Support

The evolution of learning analytics can be understood as:

Descriptive
"What happened?"
↓
Diagnostic
"Why might it have happened?"
↓
Predictive
"What may happen next?"
↓
Prescriptive
"What intervention may help?"

The most valuable systems should ultimately connect analytics with evidence-based educational interventions.

54. Human-Centred Learning Analytics

The purpose of learning analytics should remain educational improvement.

A human-centred approach prioritises:

- Student agency.
- Teacher judgement.
- Transparency.
- Privacy.
- Fairness.
- Learning outcomes.

AI should function as a decision-support mechanism rather than an autonomous authority over students.

55. Conclusion

Intelligent learning analytics represents an important development in data-driven education. By combining multimodal data with artificial intelligence, educational institutions can potentially develop richer insights into student engagement, learning behaviour and academic progress.

Behavioural, academic, textual, social and other data modalities can provide complementary information. When responsibly integrated, these sources can support early identification of learning difficulties, personalised learning, adaptive feedback and evidence-based instructional decision-making.

However, intelligent analytics also introduces substantial challenges. Student data can be sensitive, multimodal systems can generate biased interpretations, and observable behaviour does not necessarily represent genuine learning. Excessive monitoring may also create surveillance concerns and undermine student autonomy.

The proposed framework therefore emphasises multimodal data integration, strong governance, AI-based analysis, interpretable learner profiling, human-centred intervention and continuous evaluation. The future of learning analytics should not be defined by the quantity of data collected but by the quality of educational decisions enabled by that data.

Ultimately, intelligent learning analytics should help educators understand learners more effectively, help students understand their own learning processes and enable institutions to design more responsive educational environments. The most successful systems will be those that combine technological intelligence with pedagogical expertise, ethical responsibility and genuine respect for learner agency.

References

1. Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. EDUCAUSE Review, 46(5), 30–40.
2. Ferguson, R. (2012). The state of learning analytics in 2012: A review and future challenges. Technical Report, Knowledge Media Institute.
3. Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. EDUCAUSE Review, 46(5), 30–40.
4. Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. Larusson & B. White (Eds.), Learning Analytics: From Research to Practice. Springer.
5. Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1355.
6. Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. Computers in Human Behavior, 89, 98–110.
7. Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. TechTrends, 59, 64–71.
8. Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education. Educational Technology Research and Development, 68, 1967–1990.
9. Siemens, G. (2013). Learning analytics: The emergence of a discipline. American Behavioral Scientist, 57(10), 1380–1400.
10. Williamson, B. (2017). Big Data in Education: The Digital Future of Learning, Policy and Practice. SAGE.

11. Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 57(10), 1510–1529.
12. Prinsloo, P., & Slade, S. (2017). Ethics and learning analytics. In The Handbook of Learning Analytics. Society for Learning Analytics Research.
13. Drachsler, H., & Greller, W. (2016). Privacy and analytics: It's a DELICATE issue. A checklist for trusted learning analytics. Proceedings of the 6th International Conference on Learning Analytics & Knowledge, 89–92.
14. Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. British Journal of Educational Technology, 45(3), 438–450.
15. Ferguson, R. (2019). Ethical challenges for learning analytics. Journal of Learning Analytics, 6(3), 25–30.