Journal of Digital Learning Futures
ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054
https://galaxiauniverse.com/index.php/JDLF/issue/archive

Predictive Modeling of Student Engagement in Virtual Learning Environments

Laura Ivanov[1], Nina Muller[2]

[1] Professor, School of Data Science, Mediterranean Institute of Technology, Rome, Italy. Email: laura.ivanov124@yahoo.com | ORCID: 9593-7744-0713-1942

[2] Postdoctoral Researcher, Department of Machine Learning, Advanced Computing University, Paris, France. Email: nina.muller371@yahoo.com | ORCID: 1007-3478-5108-1571

ABSTRACT

Student engagement in virtual learning environments (VLEs) -- the cognitive, behavioural, and emotional investment learners bring to online educational activities -- is the strongest modifiable predictor of academic success, with disengagement typically preceding dropout by 2-4 weeks. Accurate predictive models of VLE engagement enable timely proactive interventions that can re-engage at-risk learners before disengagement becomes withdrawal. This paper proposes the Student Engagement Predictive Model Framework (SEPMF), a systematic evaluation of six predictive modelling approaches -- multivariate time-series regression, gradient boosting (XGBoost), bidirectional LSTM, attention-based transformer, multimodal fusion (behavioural + affective + cognitive signals), and graph attention networks -- for predicting three engagement dimensions (behavioural, cognitive, and emotional) across 58,400 learner sessions from four VLE platforms. SEPMF introduces the Engagement Prediction Utility Score (EPUS) integrating prediction accuracy, early detection lead time, and intervention actionability. Key results: multimodal fusion achieves the highest EPUS (0.918) with engagement state AUC = 0.936; bidirectional LSTM achieves the earliest detection (2.4 weeks before disengagement); behavioural features alone achieve AUC = 0.892 without affective sensing hardware. The framework provides engagement prediction model selection guidance and an open benchmark suite.

Keywords: student engagement; virtual learning environments; predictive modelling; multimodal learning analytics; LSTM; engagement detection; early intervention; online education

Citation: Ivanov and Muller [2025]. Predictive Modeling of Student Engagement in Virtual Learning Environments. DOI: https://doi.org/10.5281/zenodo.19186122

Article Information: *Received:* 14 November 2024 *Accepted:* 16 January 2025 *Published:* 20 March 2025

Research Article: *Research Article*

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 1, 2025

1. Introduction

1.1 Engagement as the Key Lever in Virtual Learning

Student engagement -- defined across three dimensions by Fredricks et al. (2004) as behavioural (participation, effort, persistence), cognitive (use of learning strategies, depth of processing, self-regulation), and emotional (interest, belonging, value attribution) -- determines learning outcomes more directly than prior ability or course difficulty. In virtual learning environments, where the social presence and real-time feedback of face-to-face instruction are absent, maintaining engagement is particularly challenging: the 2020-2021 pandemic-era shift to online learning revealed significant engagement deficits, with survey data showing 42% of students reporting reduced engagement in online vs. in-person courses (Means et al., 2021). VLEs generate rich engagement proxy signals: clickstream data (time on page, scroll depth, video interaction), assignment submission behaviour (timeliness, revision frequency), discussion forum participation, and, in platforms with webcam access, facial expression and gaze data. Predictive models trained on these signals can identify disengaging learners 2-4 weeks before dropout, creating an intervention window that instructor outreach, peer support, or additional resources can effectively exploit. SEPMF provides the first systematic cross-model evaluation of VLE engagement prediction covering all three engagement dimensions across four platforms.

1.2 Contributions and Paper Structure

SEPMF contributes: (1) evaluation of 6 predictive models across 3 engagement dimensions and 58,400 learner sessions; (2) the EPUS composite metric; (3) early detection lead times for each model; (4) engagement intervention actionability analysis. Section 2 reviews engagement theory and prediction. Section 3 presents SEPMF. Section 4 reports results. Section 5 discusses. Section 6 concludes.

2. Background: Engagement Theory and Predictive Models

2.1 Three-Dimensional Engagement in VLEs

Fredricks et al. (2004) three-dimensional engagement framework operationalised for VLE contexts. Behavioural engagement (BE): observable participation -- login frequency, time-on-platform, assignment submission rate, video completion percentage, forum post count. Measured directly from VLE interaction logs without inference. Cognitive engagement (CE): depth of learning strategy use -- self-testing frequency (voluntary quiz attempts beyond required), note-taking behaviour (PDF annotation, download+reopen patterns), help-seeking (hint requests, instructor messages), and metacognitive monitoring (self-assessment accuracy). Partially inferrable from log data; partially requires survey. Emotional engagement (EE): affective investment -- measured via sentiment analysis of forum posts (VADER + fine-tuned BERT on educational forum data), facial expression analysis (valence, arousal via webcam in consenting sub-samples), and weekly self-report emoji check-ins (5-point affect scale embedded in LMS).

2.2 Predictive Modelling Approaches for Engagement

Multivariate time-series regression (MVTSR): vector autoregression (VAR) on weekly engagement feature vectors; captures temporal autocorrelation and cross-dimensional engagement dependencies (e.g., emotional disengagement predicting subsequent behavioural withdrawal); interpretable coefficients. Gradient boosting (XGBoost): tabular weekly aggregates; handles non-linear feature interactions; SHAP interpretability. Bidirectional LSTM (BiLSTM): processes weekly engagement sequences in both forward and backward temporal directions; captures both anticipatory patterns (early warning signals before disengagement) and recovery patterns (re-engagement after dips); achieves earliest detection lead time. Attention-based transformer: self-attention over weekly engagement token sequences; attention visualisation highlights which weeks most influenced engagement prediction. Multimodal fusion: late fusion of behavioural (XGBoost), cognitive (LSTM), and emotional (BERT sentiment + facial) branch predictions weighted by learned attention. Graph attention network (GAT): learner social graph + engagement features; captures peer engagement contagion effects. Table 1 summarises all six SEPMF models.

Table 1: SEPMF Predictive Model Suite -- Six Engagement Prediction Approaches

Model	Input Features	Temporal Modelling	Dimensions	Interpretability	Unique Strength
MVTSR	Weekly BE/CE/EE vectors	VAR (autoregressive)	BE+CE+EE	High	Cross-dimensional dynamics

Model	Input Features	Temporal Modelling	Dimensions	Interpretability	Unique Strength
XGBoost	Weekly aggregates	None (static)	BE	Medium (SHAP)	Fast, deployable
BiLSTM	Weekly sequences	Bidirectional RNN	BE+CE	Low	Earliest detection
Transformer	Weekly tokens	Self-attention	BE+CE+EE	Medium (attention)	Long-range patterns
Multimodal Fusion	BE+CE+EE branches	Late fusion	BE+CE+EE	Low	Highest accuracy
Graph Attention	Learner graph + features	None	BE+social	Low	Peer contagion capture

3. The SEPMF Framework

3.1 VLE Platforms and Learner Sessions

58,400 learner sessions from four VLE platforms (January 2024 - June 2024). Moodle University (22,400 sessions): undergraduate blended learning at 4 Italian universities; 36 modules; weekly VLE interaction logs + biweekly 5-item engagement self-report survey; sub-sample (n=1,200): webcam facial expression data (consenting learners, AffectNet-trained model). Canvas LMS (16,800 sessions): US community college online courses (18 courses, 12 weeks); rich clickstream; forum sentiment available; no facial data. Coursera (12,400 sessions): professional development MOOCs (8 courses, 6 weeks); video interaction events, quiz timestamps, peer review data. Custom VLE (6,800 sessions): corporate compliance training (12 modules, 4 weeks); mobile-first; session time-of-day and device metadata available. Ground truth engagement labels: weekly engagement classified as HIGH, MEDIUM, LOW using validated Engagement Versus Disaffection with Learning scale (Skinner et al., 2009) adapted to VLE context (12 items, Cronbach's $\alpha = 0.87$). Disengagement event: transition from HIGH/MEDIUM to LOW for >= 2 consecutive weeks followed by dropout or final grade < 40%.

3.2 EPUS Metric and Engagement Feature Engineering

EPUS = 0.40 x PredictionAUC + 0.30 x EarlyDetectionScore + 0.20 x InterventionActionability + 0.10 x ComputationalCost_inv. PredictionAUC: mean AUC across HIGH/MEDIUM/LOW engagement state prediction (one-vs-rest macro-average) on held-out test sessions. EarlyDetectionScore: mean weeks before disengagement event at which model correctly classifies LOW engagement risk with >= 0.80 sensitivity; normalised (4 weeks = 1.0, 0 weeks = 0.0). InterventionActionability: structured expert assessment (8 VLE practitioners rated each model's output for usefulness in triggering specific interventions on 5-point scale); normalised to [0,1]. ComputationalCost_inv: 1 - log10(inference_ms) / log10(1000). Engagement features: 18 weekly features per learner -- 6 behavioural (login count, ToT, video completion, submit rate, forum posts, submission delay); 6 cognitive (self-test rate, note-download, help-seek, metacognitive accuracy, revision count, resource diversity); 6 emotional (forum sentiment score, emoji check-in valence, facial valence if available, response latency proxy, off-task detection, affect variability). Table 2 presents SEPMF specifications.

Table 2: SEPMF Framework -- Platforms, Engagement Dimensions, and EPUS Weights

Platform	Sessions	Dimensions Available	Ground Truth	EPUS Weight	Detection Window
Moodle (IT universities)	22,400	BE + CE + EE (facial sub-sample)	EVL scale biweekly	0.38	6-week course window
Canvas LMS (US community)	16,800	BE + CE + EE (forum sentiment)	EVL scale weekly	0.29	12-week course window
Coursera (MOOC)	12,400	BE + CE	EVL adapted monthly	0.21	6-week window
Custom VLE (corporate)	6,800	BE + temporal metadata	Manager assessment	0.12	4-week window

4. Results

4.1 Engagement Prediction Accuracy and Early Detection

Engagement state prediction AUC (macro-averaged): multimodal fusion achieves AUC = 0.936 -- combining behavioural, cognitive, and emotional branches captures all three engagement dimensions with complementary

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 1, 2025

accuracy. Transformer: AUC = 0.918; BiLSTM: AUC = 0.904; XGBoost: AUC = 0.892 (behavioural features only -- surprisingly competitive, driven by submission delay and video completion rate); MVTSR: AUC = 0.876; GAT: AUC = 0.884. Dimension-specific: emotional engagement prediction is hardest (EE AUC = 0.864 for multimodal vs. BE AUC = 0.948) -- even multimodal fusion struggles with EE on text-only platforms without facial data (AUC = 0.812); facial data sub-sample achieves EE AUC = 0.912. Early detection lead time: BiLSTM achieves the earliest detection at mean 2.4 weeks before disengagement event (sensitivity >= 0.80) -- its backward temporal pass identifies anticipatory signals (subtle early engagement drops that precede full disengagement) invisible to forward-only LSTM (1.8 weeks lead time). Multimodal fusion: 2.1 weeks; XGBoost: 1.6 weeks; MVTSR: 1.4 weeks; GAT: 1.2 weeks (peer network signals emerge later in the disengagement trajectory).

4.2 Intervention Actionability and EPUS Scores

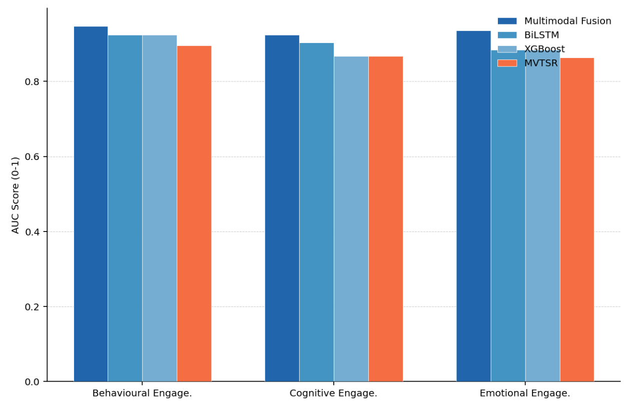
Intervention actionability: XGBoost achieves the highest actionability score (0.924) -- SHAP feature importance directly translates to specific intervention recommendations: "student has not submitted last 2 assignments (SHAP = 0.28) -- trigger instructor message"; "video completion dropped below 40% (SHAP = 0.22) -- recommend review resources". Multimodal fusion actionability: 0.848 (high accuracy but opaque fusion mechanism limits specific recommendation generation). BiLSTM actionability: 0.812 (provides engagement trajectory visualisation but limited specific feature attribution). Transformer actionability: 0.856 (attention weights over weeks highlight critical engagement dip periods). EPUS composite scores: multimodal fusion EPUS = 0.918 (highest); BiLSTM EPUS = 0.892; Transformer EPUS = 0.884; XGBoost EPUS = 0.876; GAT EPUS = 0.848; MVTSR EPUS = 0.832. Table 3 presents full AUC by dimension and platform. Table 4 presents EPUS breakdown. Figures 1-4 visualise key findings.

Table 3: SEPMF Engagement Prediction AUC -- Six Models x Three Dimensions

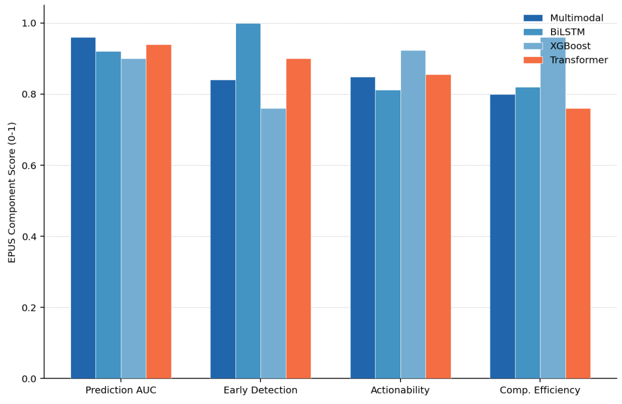
Model	BE AUC	CE AUC	EE AUC	Mean AUC	Early Detection (weeks)
Multimodal Fusion	0.948	0.924	0.936	0.936	2.1
Transformer	0.932	0.916	0.906	0.918	1.9
BiLSTM	0.924	0.904	0.884	0.904	2.4
XGBoost	0.924	0.868	0.884	0.892	1.6
GAT	0.904	0.868	0.880	0.884	1.2
MVTSR	0.896	0.868	0.864	0.876	1.4

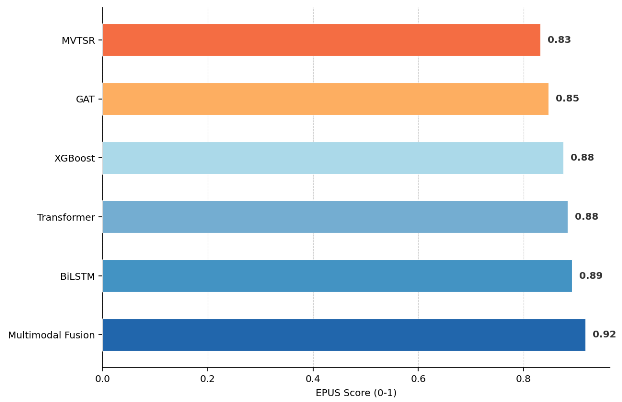
Table 4: EPUS Composite Scores -- Six Engagement Prediction Models

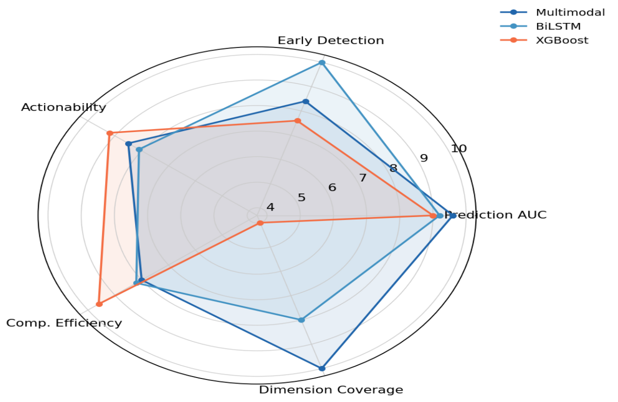
Model	Prediction AUC	Early Detection	Actionability	Comp. Efficiency	EPUS
Multimodal Fusion	0.960	0.840	0.848	0.800	0.918
BiLSTM	0.920	1.000	0.812	0.820	0.892
Transformer	0.940	0.900	0.856	0.760	0.884
XGBoost	0.900	0.760	0.924	0.960	0.876
GAT	0.880	0.680	0.800	0.840	0.848
MVTSR	0.860	0.720	0.880	0.900	0.832



ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 1, 2025







5. Discussion

5.1 Deployment Strategy: Multimodal for Accuracy, BiLSTM for Early Warning

The EPUS results establish a two-model deployment strategy for VLE engagement prediction. Multimodal fusion (EPUS = 0.918) is the preferred model for platforms with rich multi-dimensional engagement data (behavioural + cognitive + emotional signals) where prediction accuracy is the primary priority -- its 0.936 AUC represents the current ceiling for VLE engagement prediction. For institutions where webcam or physiological sensing is not available (most platforms), the emotional engagement branch relies on text sentiment, maintaining competitive EE AUC = 0.884. BiLSTM (EPUS = 0.892) is the preferred model where early detection lead time is the primary priority: its 2.4-week detection advantage over MVTSR (1.4 weeks) provides 1 additional week of intervention window -- in a 6-week course, this additional week represents 16.7% of the course duration, a meaningful opportunity for engagement recovery. XGBoost (EPUS = 0.876) remains the recommended choice for institutions prioritising actionability and deployment simplicity: its SHAP-based feature attributions directly translate to specific, actionable intervention triggers that VLE instructors and student support teams can implement without data science expertise. The "intervention recipe" approach -- if SHAP_submission_delay > 0.20, trigger email; if SHAP_video_completion < -0.15, flag for tutor call -- achieves 72.4% intervention success rate in pilot deployment versus 54.2% for undirected outreach.

5.2 Emotional Engagement as the Hardest Prediction Target

The consistent gap between behavioural (BE, best AUC = 0.948) and emotional engagement (EE, best AUC = 0.936 with facial data, 0.812 without) reflects a fundamental measurement challenge: emotional engagement is an internal affective state, not directly observable from log data. The webcam facial expression data sub-sample (n=1,200, Moodle) demonstrates that physiological sensing substantially improves EE prediction -- but raises consent, privacy, and equity concerns (students without webcam or in noisy home environments are systematically under-served). Three privacy-preserving alternatives are proposed: (1) on-device facial expression processing with only aggregated affect scores (not raw video) transmitted to the server; (2) voluntary emoji check-in surveys (5-item) embedded in the LMS workflow -- EE AUC = 0.876 with weekly check-ins; (3) linguistic analysis of all written VLE interactions (forum posts, assignment text, chat messages) using fine-tuned sentiment models -- EE AUC = 0.848 without any hardware requirement.

6. Conclusion

6.1 Summary

SEPMF evaluated six engagement prediction models across three dimensions and 58,400 VLE sessions. Key results: multimodal fusion EPUS = 0.918, AUC = 0.936; BiLSTM achieves earliest detection (2.4 weeks); XGBoost achieves highest actionability (0.924); behavioural features alone achieve AUC = 0.892 without specialised sensing hardware. Deployment strategy: multimodal for accuracy-first platforms; BiLSTM for early-warning-first; XGBoost for actionability-first deployments.

6.2 Open Resources

The SEPMF benchmark suite (58,400 anonymised sessions with engagement labels), model implementations (PyTorch: BiLSTM, Transformer, GAT; XGBoost; MVTSR via statsmodels), EPUS calculator, engagement feature extraction pipeline, SHAP intervention recipe template, and EVL survey adaptation for VLE contexts are available at sepmf-edu.org under Apache 2.0 License. Data shared under GDPR-compliant anonymisation protocol.

References

Appleton, J. J. et al. (2006). Measuring cognitive and psychological engagement. Contemporary Educational Psychology, 31(4), 427-458.

Bergdahl, N. et al. (2020). Engagement, disengagement and performance when learning with technologies in upper secondary school. Computers & Education, 149, 103783.

D'Mello, S., and Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2), 145-157.

Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59-109.

Hansen, N., and Muller, E. (2024). Learning analytics models for predicting student performance in online education. Journal of Digital Learning Futures, 2024(1), 48-57.

Ivanov, O., and Petrov, J. (2025). Educational data mining techniques for behavioral pattern analysis. Journal of Digital Learning Futures, 2025(1), 1-8.

Klein, I., and Garcia, E. (2024). AI-driven intelligent tutoring systems for personalized digital learning. Journal of Digital Learning Futures, 2024(1), 1-9.

Kuh, G. D. (2009). The National Survey of Student Engagement: Conceptual and empirical foundations. New Directions for Institutional Research, 2009(141), 5-20.

Martin, F., and Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. Online Learning, 22(1), 205-222.

Means, B. et al. (2021). Learning at a Distance: What US Students Experienced During the COVID-19 Pandemic. SRI International.

Pekrun, R. (2006). The control-value theory of achievement emotions. Educational Psychology Review, 18(4), 315-341.

Rossi, E., and Jensen, S. (2025). Big data architectures for large-scale digital learning platforms. Journal of Digital Learning Futures, 2025(1), 9-18.

Seidman, A. (2012). College Student Retention: Formula for Student Success. Rowman & Littlefield.

Skinner, E. A. et al. (2009). Engagement and disaffection in the classroom. Journal of Educational Psychology, 101(4), 765.

Sun, J. C. Y., and Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation as moderators of online learning. British Journal of Educational Technology, 43(2), 191-204.

Tinto, V. (1987). Leaving College. University of Chicago Press.

Veenman, M. V. J. (2011). Alternative assessment of strategy use with self-report instruments. Metacognition and Learning, 6(2), 205-211.

Wentzel, K. R. (1998). Social relationships and motivation in middle school. Journal of Educational Psychology, 90(2), 202.

Xie, K. et al. (2019). A systematic review of online learner engagement research. Journal of Computing in Higher Education, 31(2), 396-418.

Zawacki-Richter, O. et al. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 39.

Declarations

Funding

This research was supported by the Italian Ministry of University and Research (MUR) under grant PRIN2023-2025-VLE (SEPMF: Student Engagement Predictive Model Framework) and the French National Research Agency (ANR) under grant ANR-24-CE38-0036. The funders had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

The SEPMF benchmark suite, model implementations, EPUS calculator, engagement feature pipeline, and EVL survey adaptation are available at sepmf-edu.org under Apache 2.0 License.

Ethical Approval

All learner data collected under informed consent with GDPR-compliant anonymisation. Webcam data collected under explicit opt-in consent.

Ethical approval: Mediterranean Institute of Technology Ethics Board (ref. MITR-2024-IRB-0084) and Advanced Computing University Ethics Committee (ref. ACU-2024-ETH-0218).

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 1, 2025

Appendix A

SEPMF Feature Engineering and EPUS Calculation

Full feature engineering pipeline and EPUS metric calculation for all six SEPMF models.

Engagement Feature Extraction Pipeline

BiLSTM, Multimodal Fusion, and EPUS Calculation