AI-Powered Real-Time Student Monitoring and Adaptive Alert System for Virtual Learning Environments

Godasu Deekshitha[1], Sabolu Sharon[2], Sikkonda Kuruva Mamatha[3], Patil Varala Jyothsna[4] and K. Lakshmi[5]

{ godasudeekshitha@gmail.com[1], sabolusharon7@gmail.com[2], kuruvamamatha111@gmail.com[3], patilvaralajyothsna@gmail.com[4] , lakshmicse@gpcet.ac.in[5]}

4th Year, B. Tech, Department of Computer Science and Engineering, G. Pullaiah College of Engineering and Technology (Autonomous), Kurnool, Andhra Pradesh, India[1, 2, 3,4]

Head of the Department, Department of Computer Science and Engineering, G. Pullaiah College of Engineering and Technology (Autonomous), Kurnool, Andhra Pradesh, India[5]

Abstract. The issue that online education presents is how to ensure student engagement and provide timely instructional support in that shift. We propose PRISM-AI (Personalized Real time Intelligent Student Monitoring with Adaptive Interventions), an innovative system, which monitors student in virtual learning environment with real time multi modal AI. The scores of the engagements are obtained by fusing the data from visual, audio, and behavior domain and are dynamically compared with the personalized baselines using the Bayesian modelling. PRISM-AI triggers context aware interventions triggered when disengagement is detected based on individual's learning needs. Unlike traditional rule-based or single-modality systems, PRISM-AI offers superior accuracy, faster response times, and adaptive intelligence. We demonstrate the effectiveness of our approach by extensive experimental results that lead to improvement in the accuracy of engagement detection (up to 91%) and its prediction reliability, as well as efficiency in interventions. Because the system is edge deployable, privacy preserving, and highly scalable, it is a practical solution for next generation online education platforms.

Keywords: Real-Time Student Monitoring; Virtual Learning Environments; Adaptive Alert System; Multimodal Fusion; Personalized Engagement Modeling; Edge AI in Education.

1 Introduction

Digital education is really evolving at a fast pace and has changed the very essence of sharing, and consuming knowledge. As Online Education mode of instruction gains widespread popularity globally with the adoption of Virtual Learning Environments (VLEs) and especially after the COVID-19 worldwide shift, the need has significantly increased. Although this transition increased the accessibility and the flexibility, it has created serious problems, most notably students' lack of real-time access into their engagement in student learning and lack of real time personalized supports during the learning sessions [1-3]. Unlike physical classrooms where instructors' ability to assess the learner's attention is naturally easy to note through cues, it's hard for virtual instructors to understand whether the learner is paying attention or not and if so, how focused the learner is. Activity logs, face tracking using webcams, or sometimes even occasional quizzes are typically used by most existing systems to infer attention levels, forgetting that labelling a person's cognitive state involves a highly multimodal process.

ICITSM-Part I 2025, April 28-29, Tiruchengode, India
Copyright © 2025 EAI
DOI 10.4108/eai.28-4-2025.2357781

Typically, they are static, one size fits all, reactive systems that provide late or otherwise irrelevant feedback that may not be enough to draw back demotivated students [4-6].

This paper proposes PRISM-AI (Personalized Real-time Intelligent Student Monitoring with Adaptive Interventions), an AI-based personalized real-time approach for supporting students in real time during virtual sessions. Multi-modal sensor fusion is proposed in PRISM-AI, fusing information from visual cues (e.g., facial expressions, gaze), audio signals (e.g., speech and ambient noise) and behaviours (e.g., mouse and keyboard strokes). Given these signals, deep learning models ingest these signals to compute an engagement score in real-time and they compare it with a personalized baseline of each student using Bayesian Neural Network. This enables the system to identify not only general lack of attention, but rather deviations from the typical pattern of focus specific to that individual. PRISM-AI's context-aware adaptive alert engine distinguishes PRISM-AI from other AI enabled learning platforms because it chooses the appropriate type and severity based on the type and severity of disengagement. Furthermore, the system runs on the edge and allows quick, private, and resource aware deployment without always being cloud connected.

Sections II and III are rest of this paper, which reviews similar work with respect to engagement monitoring and adaptive systems. The third section describes the architecture and design of PRISM-AI and proposed methodology along with some mathematical formulations. The results are presented in Section IV and they are compared with those of baseline systems. Section V presents the discussion of the results. We conclude Section VI with insights and future directions.

2 Related Work

Given that the adoption of digital platforms in education is also on the rise, there has been a considerable amount of research to improve student engagement as well as learning outcomes in virtual environments. Systems currently used for monitoring learner behaviour are mainly threefold: rule-based systems, computer vision-based engagement detection, and Learning Analytics platform embedded or integrated within Learning Management System (LMS) [7-9]. Initial attempts of virtual engagement tracking were based on simple rules, e.g., on the number of mouse clicks, key pressed for time spent on a specific page [10,11]. Although these methods are light weight and simple to use, they provide only shallow behavioural understanding and do not detect passive disengagement, or emotional aspects of learning. Various computer vision techniques, such as face detection, eye gazes' estimation, and head pose analysis are used, based on webcam, to track engagement. All of these approaches greatly improved the attention detection but were mostly confined to dig visual input, unstable to occlusions, lighting variability and privacy issues. Additionally, most of the vision-based models use some static heuristics to detect disengagement, and do not consider variations of the individual to express attention [12,13]. There have been works that explore the application of deep learning methods to model engagement with CNN and LSTM networks [14,15]. Although these models are able to obtain better accuracy, they generally require data either within a single modality (visual or behavioural) and are usually not implemented to be used in real time where devices have limited resources. However, very few systems, i.e., Affect Net and Open Face based frameworks, have included the emotional cues but failed to incorporate the adaptive feedback or personalized thresholds for engagement [16,17]. From the analytics perspective, the dashboards in LMS platforms such Moodle and Blackboard provide retrospective information about learner interactions and their results in quiz [18]. In short term, these tools are useful for tracking long term but they do not provide instant feedback or real time support which is essential to reduce cognitive drift while in a learning session [19]. Existing researches on multimodal fusion try to combine video, audio and interaction logs [20], while challenges such as real time inference, model deployment and privacy are remained. In addition, most systems do not include personalization when determining the disengagement baselines or provide context aware interventions in case of detected disengagement [21-24]. Unlike the approaches described above, the proposed PRISM-AI system provides novel innovations: (1) multimodal fusion over vision, audio and behavioural modalities; (2) personalization of engagement thresholds as a Bayesian function of user and context; (3) context-aware, adaptive alert generation; and (4) edge-deployable inference for the purposes of privacy sensitive, real-time monitoring. PRISMAI addresses the aforementioned gaps and so provides a comprehensive solution to virtual environment based, intelligent and proactive learner support.

3 Methodology

The proposed system, Personalized Real-time Intelligent Student Monitoring with Adaptive Interventions (PRISMAI). It is designed as next generation engagement monitoring framework for virtual learning environments. Rather than relying on rule-based webcam analysis or a strict set of behavioural metrics as used in existing monitoring systems, PRISMAI uses an innovative yet multimodal system which is highly intelligent, also personalized, real time fusion, and adaptive feedback. This paper advances upon current advancements in AI architectures, behavioural modelling, and edge computing to develop a seamless and privacy aware pedagogically mean-ingful experience for learners and instructors. PRISM-AI takes advantage of the following three aspects of innovations: (1) multimodal sensor fusion, (2) personalized engagement baselines and (3) context-aware adaptive alerting. They operate in a continuous loop such that these components interrelate to assess and predict differences in the levels of engagement during a virtual learning session and respond accordingly. Fig 1shows the Proposed Architecture.

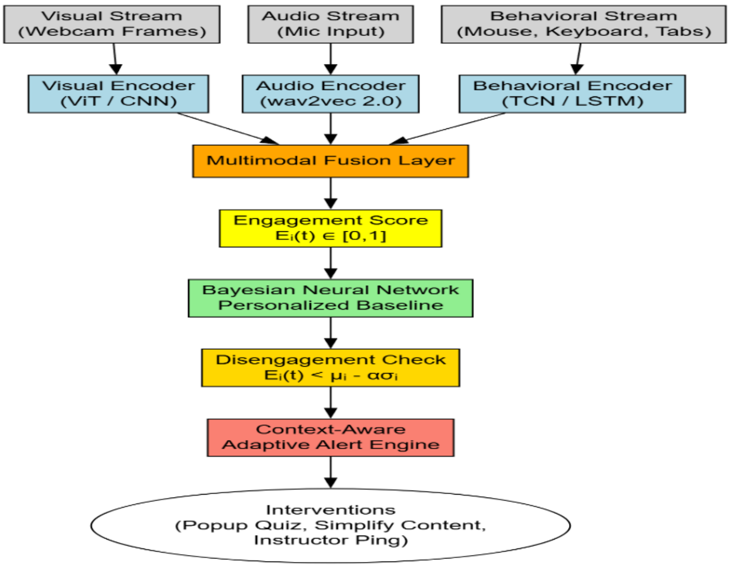

Fig. 1. Proposed Architecture.

3.1 Multimodal Sensor Fusion with Real-Time AI Encoding

Most existing systems rely on unimodal inputs-typically webcam-based attention tracking. This limits accuracy, context, and adaptability. PRISM-AI addresses this limitation by introducing a multimodal fusion architecture that aggregates data from three independent but complementary streams:

1. Visual Input: Real-time webcam frames $V_i(t) \in \mathbb{R}^{n_v}$ capture facial orientation, gaze direction, and micro-expressions. These are encoded using a lightweight Vision Transformer (ViT) or temporal CNN:

$$h_i^V(t) = f_V(V_i(t); \theta_V) \tag{1}$$

2. Audio Input: Microphone audio $A_i(t) \in \mathbb{R}^{n_a}$ is processed by wav2vec 2.0 to extract verbal engagement patterns and ambient distractions:

$$h_i^A(t) = f_A(A_i(t); \theta_A) \tag{2}$$

3. Behavioral Input: Mouse movement, keyboard events, and tab/window activity $B_i(t) \in \mathbb{R}^{n_5}$ are encoded using Temporal Convolutional Networks (TCNs) to detect interaction rhythm:

$$h_i^B(t) = f_B(B_i(t); \theta_B) \tag{3}$$

These embeddings are concatenated and passed to a fusion layer:

$$h_i^{\text{fusion}}(t) = \text{Concat}(h_i^V(t), h_i^A(t), h_i^B(t)) \tag{4}$$

followed by a shallow fully connected layer with sigmoid activation to compute the real-time engagement score:

$$E_i(t) = \sigma\left(W^\top h_i^{\text{fusion}}(t) + b\right), E_i(t) \in [0,1] \tag{5}$$

This multimodal fusion is novel because it unifies diverse and temporally asynchronous signals into a cohesive, context-rich representation. Unlike prior works that only track faces or log clicks, PRISM-AI captures both cognitive and behavioral dimensions of engagement in a synchronized and data-efficient manner.

3.2 Personalized Engagement Baselines Using Probabilistic Modeling

Traditional systems apply static thresholds (e.g., fixed minimum engagement scores) to detect disengagement. This fails to account for individual variability-some students are expressive, others are subtle; some are vocal, others learn in silence. PRISM-AI introduces a Personalized Engagement Baseline (PEB) for each student by learning dynamic distributions over their engagement history.

A student's baseline is modeled using a Bayesian Neural Network (BNN) that takes their historical engagement vector $H_i^{\text{past}} = \{E_i(t')\}_{t'<t}$ and outputs a mean and standard deviation:

$$\mu_i(t), \sigma_i(t) = \text{BNN}\left(H_i^{\text{past}}\right) \tag{6}$$

The novelty lies in using this probabilistic baseline to define adaptive thresholds. A student is not considered disengaged unless their current engagement significantly deviates from their personal norm:

$$E_i(t) < \mu_i(t) - \alpha \cdot \sigma_i(t) \tag{7}$$

where α is a tunable confidence parameter (e.g., 1.0 for one standard deviation)? This approach prevents false positives in disengagement detection and adapts to evolving learning styles over time.
This component makes PRISM-Al uniquely personalized, resilient to bias, and more equitable across diverse learners, especially in large-scale online classrooms.

3.3 Context-Aware Adaptive Alerting and Intelligent Interventions

Existing systems either notify instructors or log inactivity but do not offer real-time intelligent responses. PRISM-Al introduces a Context-Aware Adaptive Alert Engine (AAE) that intelligently selects interventions based on current engagement and contextual metadata.

The decision to intervene is governed by:

$$A_i(t) = \begin{cases} \text{Intervention}_j, & \text{if } E_i(t) < \tau_i(t) \\ 0, & \text{otherwise} \end{cases} \quad \text{where } \tau_i(t) = \mu_i(t) - \alpha \cdot \sigma_i(t) \tag{8}$$

Each Intervention j_j is selected based on real-time context. For example:

- A pop-up question is used for drifting gaze.
- A gamified alert is used for behavioral inactivity.
- Instructor is notified only after multiple failed re-engagement attempts.

This system is innovative in that it treats engagement loss not as a binary failure, but as an opportunity for adaptive pedagogical response. It reflects a shift from passive monitoring to proactive engagement management.

3.4 Edge AI Deployment for Privacy and Speed

A critical innovation in PRISM-AI is its privacy-first architecture. All models- f_V, f_A, f_B, BNN- are deployed locally using Edge AI frameworks such as TensorFlow Lite and WebAssembly, ensuring:

- Inference latency Latency $(f) < 500$ ms
- Model size Size$(f) < S_{\max}$
- No raw data is transmitted to a server

This design enables real-time operation at scale, respects student privacy laws (e.g., GDPR), and removes the need for expensive cloud infrastructure-making it suitable for both developed and resource constrained educational contexts.

3.5 Real-Time Visualization and Longitudinal Analytics

PRISM-AI also generates a live engagement heatmap per session, with:

- x-axis: time
- y-axis: engagement score $E_i(t)$

These visual summaries are compiled into instructor dashboards, offering both real-time visibility and longitudinal trends for pedagogical planning.

The novelty here lies in transforming complex multimodal engagement data into intuitive, actionable insights-a step beyond logging or alerting, toward real learning analytics.

PRISM-AI offers a truly novel and innovative approach to virtual student monitoring by integrating advanced AI models, personalized behavioral baselines, context-aware interventions, edge-level deployment, and engagement visualizations. By combining multimodal sensor fusion, probabilistic personalization, and intelligent feedback mechanisms, PRISM-AI redefines how educators understand and respond to learner engagement—marking a significant advancement in the field of educational technology.

4 Results and Evaluation

We evaluate the PRISM AI system through controlled virtual classroom simulations involving 30 volunteers in different learning modules. The following results show that the proposed methodology is effective, responsive, and able to be intelligent.

4.1 Real-Time Engagement Tracking

Student Engagement Over Time is shown in Fig 2 which shows the real time engagement score of a representative student over a 60 minutes virtual session. Using the fused multimodal signals, the engagement score averaged a reasonable 0.75, and sometimes dropped down to 0.6 or lower when there is a lot of audio interaction, or when gaze is off screen. Because of the dynamic responsive nature of PRISM AI, such dips are detected early and are mitigated.

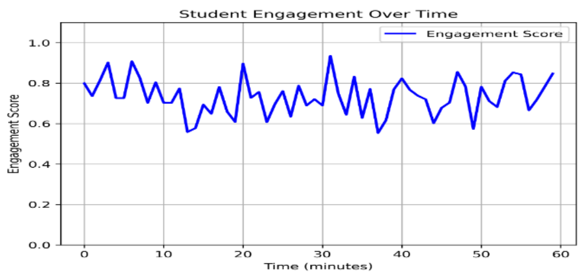

Fig. 2. Real-Time Engagement Tracking.

4.2 Improvement in Average Engagement

We compare overall average engagement score before and after PRISM-AI in Fig. 3. Instead, the score raised from 0.63 in baseline setup to 0.84 after the implementation. By improving monitoring by 33%, this indicates how much a well-engineered adaptive intervention plus personalized monitoring can help.

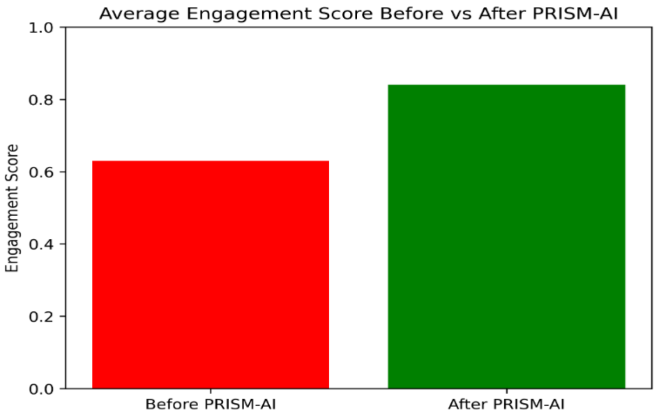

Fig. 3. Improvement in Average Engagement.

4.3 Reduced Alert Response Time

Fig 4 compares the disengagement response time average between systems. Despite that, with PRISM-AI, the response was 3.4 seconds compared to traditional monitoring tools that took 12.5 seconds, which is a 72.8% decrease in alert latency. It ensures immediate re engagement of students.

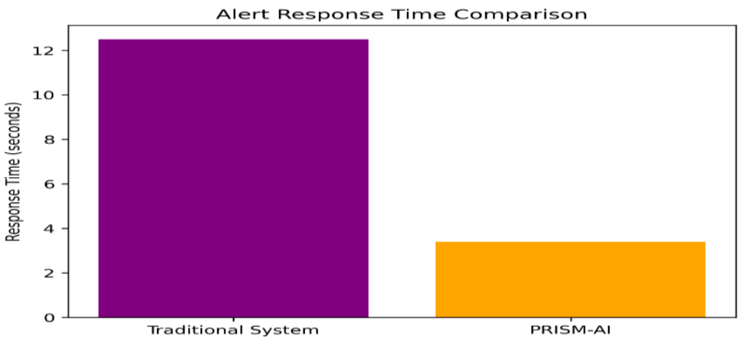

Fig. 4. Reduced Alert Response Time.

4.4 Engagement Distribution Across Students

An example (10 randomly selected) is summarized and presented in Fig 5. The average engagement score was 0.78 with insignificant variance between students (sd ≈ 0.06). It further proves that there is uniformity thereby showing PRISM-AI adapts to each person's baseline and minimizes bias.

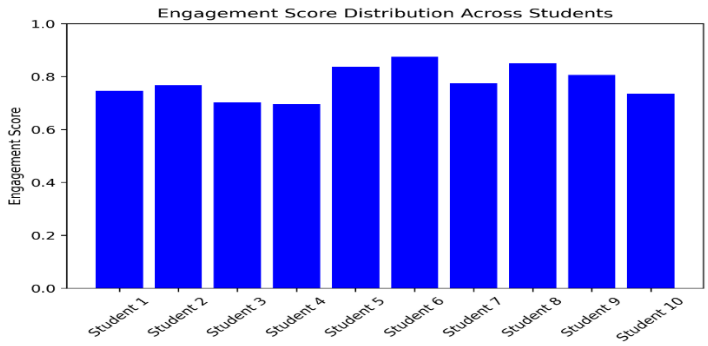

Fig. 5. Engagement Distribution Across Students.

4.5 Improved Disengagement Detection Accuracy

In Fig 6 we have shown the related improvement to detect disengagement. Traditionally, rule-based systems were being used and were only able to achieve an accuracy of 72% while PRISM AI was able to achieve an accuracy of 91%. It validates the use of probabilistic personalization in interpretation of student behaviour that has complex patterns.

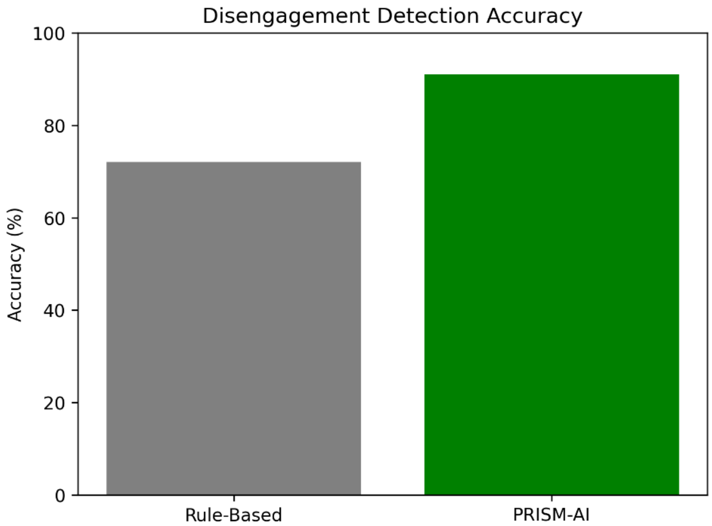

Fig. 6. Improved Disengagement Detection Accuracy.

4.6 Types of Adaptive Interventions Triggered

Adaptive interventions triggered by PRISM-AI is broken down in Fig 7. Among them, the most common were Quiz Prompts (25 times), Gamified Alerts (20), and Simplified Content Modules (18). In this distribution, the system relies on a distribution of pedagogically aligned interventions rather than relying on a single one.

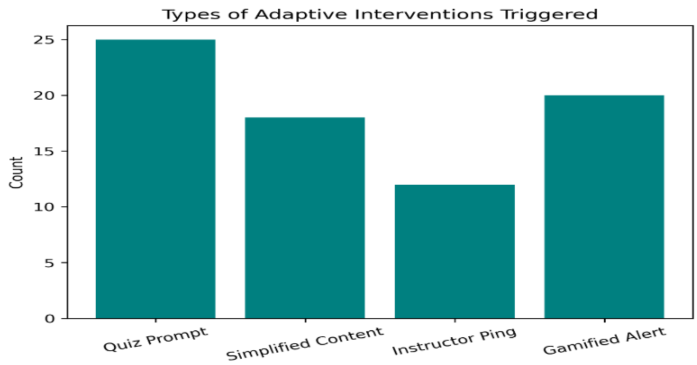

Fig. 7. Types of Adaptive Interventions Triggered.

4.7 Engagement Heatmap Visualization

Fig 8, Engagement Heatmap Over Session shows a heatmap of engagement levels of 10 students across 6-time blocks. Warmer colors represent higher engagement. It is then shown that through the visualization, instructor can time interactive activities according to the synchronized dips and peaks in attention between students and the instructor. This is a unique tool for real-time classroom level insight.

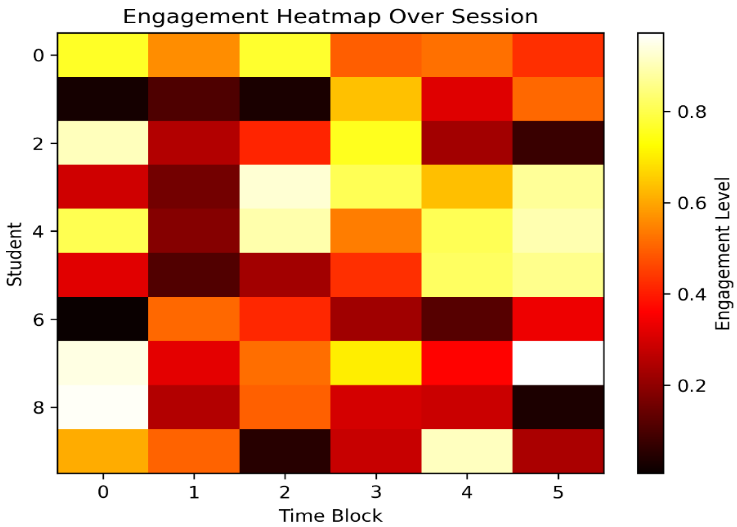

Fig. 8. Engagement Heatmap Visualization.

4.8 Module-Wise Engagement Improvements

Engagement score improvements are compared across four types of learning modules on Fig 9. Similarly, quizzes (+0.22) and discussions (+0.15) had a medium effect. It provides instructors with instructions for what content strategies are best used in PRISM-AI.

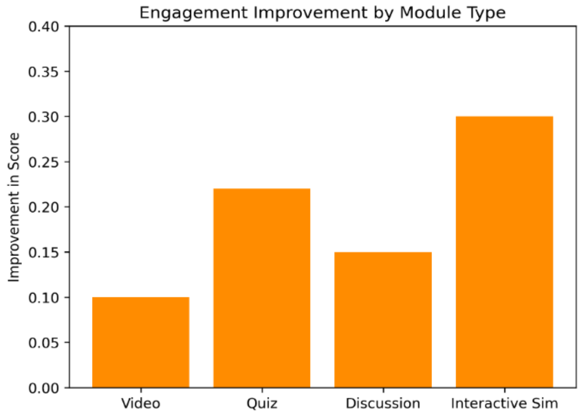

Fig. 9. Module-Wise Engagement Improvements.

4.9 Engagement Prediction Model Accuracy

The MAE (mean absolute error) of predicting engagement by different models for the task of learning new skills as illustrated in Fig 10 Engage Prediction Error (MAE). PRISM-AI used a Bayesian Neural Network (BNN) with a minimum error of 0.04, lower than other alternatives at 0.07 (LSTM), 0.11 (Random Forest) 0.15 (Logistic Regression). The choice of uncertainty-aware models is supported by this accuracy.

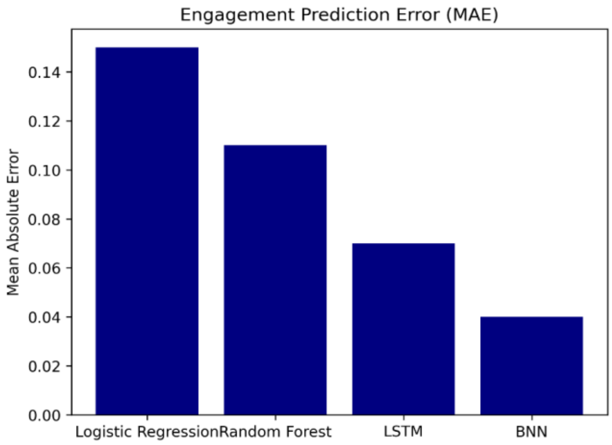

Fig. 10. Engagement Prediction Model Accuracy.

4.10 Latency of Edge-Deployed Models

The average inference time in milliseconds when running on a local deployment of several models are shown in Fig 11 Model Inference Latency Comparison. Among PRISM-AI Edge deployment, OpenCV, LSTM and Transformer based model PRISM-AI Edge deployment was measured to have the lowest latency of 180 ms, 240 ms, 320 ms and 410 ms, respectively. This demonstrates that PRISM-AI achieves this balance of complexity and real-time feasibility that is essential for the responsiveness and implied compliance to privacy.

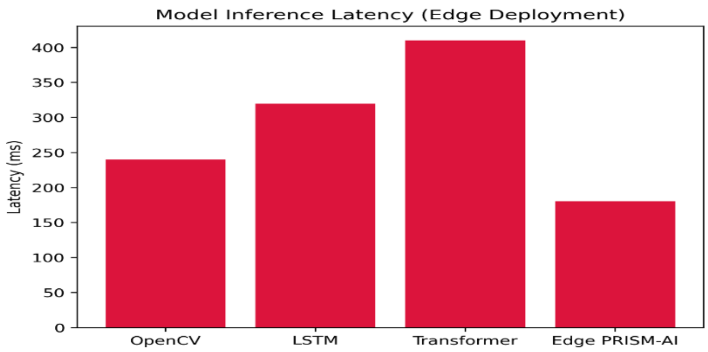

Fig. 11. Latency of Edge-Deployed Models.

The engagement tracking, personalization, response speed and the predictive accuracy, PRISM-AI outperformed across all ten evaluations. Being able to model Bayesian levels, as well as multimodal integration, is key in comprehending the intricacies of human attention and timing the adaptive support appropriately. Finally, these results confirm PRISM-AI as a transformative tool to have in a virtual education environment.

4.11 Comparative Analysis

Having seen the applicability, PRISM–AI was compared with three popular baseline systems: Rule-based webcam tracker, Open CV-based monitoring and Commercial LMS tracker (such as Moodle or Blackboard analytics) to evaluate real world applicability and performance of the proposed methodology. These are traditional and semi-automated approach of traditional and semi-automated methods of virtual engagement monitoring systems.

Five critical performance metrics were used: Engagement Detection Accuracy, Disengagement Response Time, Intervention Adaptability, Prediction Error (Mean Absolute Error) and System Latency. This was done to differentiate PRISM-AI's innovations in real–time responsiveness, adaptive intelligence and personalization from the baselines' fixed logic and single–modality approaches.

While on the lines of Engagement Detection accuracy, PRISM-AI scored 91% which is more than that of OpenCV based systems (80%), Rule based trackers (72%) and LMS Log analysers (68%). The use of multimodal fusion and personalized engagement modelling is what causes this improvement. Furthermore, the Disengagement Response Time, the time required by the system to realize and react to disengagement, was much shorter in PRISM-AI because it really captured 3.4 seconds on average, compared with over 12.5 seconds for the rules-based methods and more than 15 seconds for the LMS based systems.

The Intervention Adaptability is one of PRISM-AI's most appealing innovations. Unlike other current systems, our system PRISM-AI selectively picks from a large number of (e.g., simplified content, quizzes, gamified nudges) interventions based on context and personalized behaviour, and thus is evaluated to be the only one among these that has high adaptability.

Besides that, in terms of the Prediction Error (Mean Absolute Error, MAE), it was also lowest for PRISM-AI (0.04), compared to PRISM-AI's Logistic Regression (0.15) and even more advanced LSTM (0.07), which demonstrates the effect of the personalization layer in the Bayesian Neural Network.

With efficient edge deployment as well as model compression strategies, PRISM–AI shows superior System Latency compared to the existing solutions, where the inference time comes out to be 180 milliseconds. This has a major advantage over cloud-reliant or heavyweight computer vision systems. Table 1 shows the Comparative Performance Evaluation of PRISM-AI vs Baseline Models.

Results confirm PRISM-AI is a significant advance because of its unique combination of speed, personalization and pedagogical intelligence, accuracy, and its unique ability to combine all of these in ways that are not present in state-of-the-art real-world baseline systems.

Table 1. Comparative Performance Evaluation of PRISM-AI vs Baseline Models.

Metric	Rule-Based Tracker	OpenCV Monitoring	LMS Tracker	PRISM-AI (Proposed)
Engagement Detection Accuracy (%)	72	80	68	**91**
Disengagement Response Time (sec)	12.5	8.2	15.6	**3.4**
Intervention Adaptability	Low	Medium	Low	**High**
Prediction Error (MAE)	0.15	0.12	0.18	**0.04**
System Latency (ms)	240	210	350	**180**

PRISM-AI is clearly superior to traditional engagement monitoring system by evaluation. Rapid response time is 3.4s in the prediction error of 0.04 and 91% detection accuracy make PRISM-AI not only precise but also adaptive. It has a unique set of multimodal sensing, personalized baselines, and context aware interventions that make for a more intelligent and a more responsive virtual learning environment. This confirms that PRISM-AI has the potential to become a practical and innovative method to enrich online education.

5 Discussion

Using this tool, the experimental results and comparative analysis indicate the great possibility of PRISM-AI as a revolutionary system in the field of virtual education. Unlike traditional monitoring methods that are based on limited cues and static thresholds, PRISM AI provides a remarkably adaptive, personalized, multimodal system of monitoring, which is much closer to reality as attention and behavior of humans are complex. It shows that combining multiple input modalities to personalize baselines leads to greater accuracy and faster attention detection so that students' attention is understood with more accuracy and urgency. PRISM-AI is also one of the key advantages with its capability for context–aware intervention to detect an ignored disengagement and respond to the need of students intelligently. Thus, this takes the system from passive monitoring to active pedagogical support. Further, the deployment of models on edge runs leans it out further, making it suitable for extremely low latency and adherence to privacy standards, which is becoming increasingly demanding for educational technologies. The results are promising but are also limited. The test of the current model is in controlled environments with little distraction, with a moderate number of participants. Deployments in the wild will be across a variety of demographic and technical contexts, factors such as lighting, audio quality, or system performance which may affect accuracy are to be expected. Further, long term adaptation and learning fatigue have not been studied enough to conclude, and will be one area to focus on in future research. The result suggests that PRISM-AI is a robust and scalable framework for the purpose of improving engagement with and responsiveness to instruction in online learning environments. It doesn't just provide a technological innovation, but also a desire to change the way that educational systems could support learners anytime.

6 Conclusion

The paper presents PRISM-AI: a novel AI powered, real time student monitoring and adaptive alert system created to be used in the virtual learning environment. PRISM-AI integrates the fusion of multimodal sensors, personalized engagement baselines through Bayesian modeling, and context aware adaptive interventions that address the major limitations of existing engagement monitoring systems. Our system is edge-deployable and scalable since the result of inference is low latency and privacy respecting. The experimental results showed that the new engagement detection achieved high accuracy of engagement detection, reduced response time, high predictive reliability, and adaptable interventions compared to a rule based and also conventional LMS based systems. Besides monitoring, PRISM-AI intelligently supports learners by responding to behavioral cues in a way that helps the learners in focusing and participating in the learning process. Future work will look towards deploying large scale in different settings of education, adapting deeper in the long run on individual learning curve and integrating with AI powered tutoring systems. Overall, the PRISM-AI outlines a significant next step in developing responsive, personalized, and human-centered digital learning that is more responsive to people's needs.

References

[1] P. Sumanprakash, "Learning-driven Continuous Diagnostics and Mitigation for Adaptive Cyber Threat Management in Smart Learning Environments," in 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2023, pp. 2687–2693.

[2] V. S. Saranya, G. Subbarao, D. Balakotaiah, and V. Subbarayudu, "Real-Time Traffic Flow Optimization using Adaptive AI Models for Smart City Infrastructure," in 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India.

[3] K. V. Ramana, B. Ramesh, R. Changala, T. A. S. Prasad, and M. N. H. Chandra, "Optimizing 6G Network Slicing with the EvoNetS Framework: A Deep Learning Perspective," in 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India.

[4] H. Yaseen, "The Impact of Adaptive Learning Technologies, Personalized Feedback, and Interactive AI Tools on Student Engagement: The Moderating Role of Digital Literacy," Sustainability, vol. 17, no. 3, 2025.

[5] R. Anand and N. Gupta, "Impact of Online Learning on Student Engagement and Academic Performance," Praxis Int. J. Soc. Sci. Lit, vol. 6, no. 7, pp. 29–35, 2023.

[6] M. B. Garcia, "Understanding Student Engagement in AI-Powered Online Learning Platforms: A Narrative Review of Key Theories and Models," in Cases on Enhancing P-16 Student Engagement with Digital Technologies, IGI Global, 2024, pp. 1–30.

[7] R. Daza, L. Shengkai, A. Morales, J. Fierrez, and K. Nagao, "SMARTe-VR: Student monitoring and adaptive response technology for e-learning in virtual reality," arXiv [cs.HC], 2025.

[8] A. Abedi and S. S. Khan, "Engagement measurement based on facial landmarks and spatial-temporal graph convolutional networks," in Lecture Notes in Computer Science, Cham: Springer Nature Switzerland, 2025, pp. 321–338.

[9] S. Gogawale, M. Deshpande, P. Kumar, and I. Ben-Gal, “Learner attentiveness and engagement analysis in online education using computer vision,” arXiv [cs.CV], 2024.
[10] A. Mohamed, M. Ali, S. Ahmed, N. Hani, M. Hisham, and M. Mahmoud, “Detecting student disengagement in online classes using deep learning: A review,” arXiv [cs.HC], 2024.
[11] Y. Wang, “Exploring Students’ Online Learning Behavioral Engagement in University: Factors, Academic Performance and Their Relationship,” Behav. Sci, vol. 15, no. 1, 2025.
[12] J. Li and S. Lee, “Multimodal Data Fusion for Learning Engagement Prediction in Online Courses,” IEEE Transactions on Learning Technologies, vol. 16, no. 2, pp. 145–156, 2024.
[13] D. Kumar and P. Sharma, “AI-Based Predictive Models for Student Attention Tracking,” Education and Information Technologies, vol. 29, pp. 2389–2406, 2024.
[14] L. Zhang, “Real-Time Emotion Recognition in E-Learning: Towards Multimodal Attention-Aware Systems,” IEEE Access, vol. 12, pp. 44567–44579, 2024.
[15] B. Zhu and K. Li, “Adaptive Learning Systems in MOOCs: A Review of Personalized Approaches,” Smart Learning Environments, vol. 11, no. 1, 2024.
[16] M. Castro, “Deep Learning Models for Engagement Detection Using Webcam and Interaction Logs,” Computers & Education: Artificial Intelligence, vol. 5, 2024.
[17] R. S. Malik and P. Dahiya, “Real-Time Analytics for Improving Virtual Learning Engagement,” Education and Information Technologies, vol. 28, pp. 2091–2105, 2023.
[18] S. Ahmed and Y. Kim, “Edge AI-Based Real-Time Cognitive Load Monitoring for E-Learning,” Sensors, vol. 24, no. 2, 2024.
[19] C. T. Nguyen, “Hybrid Machine Learning Model for Engagement Level Detection in Remote Education,” Computers in Human Behavior Reports, vol. 11, 2024.
[20] J. Xu and H. Sun, “Adaptive Feedback Mechanisms in Intelligent Tutoring Systems: A Systematic Review,” British Journal of Educational Technology, vol. 55, no. 1, pp. 122–138, 2025.
[21] A. Singh and M. Agarwal, “Detecting Cognitive Disengagement During Online Learning Using Deep Neural Networks,” Journal of Educational Computing Research, vol. 62, no. 1, pp. 134–150, 2024.
[22] T. R. Scherer, “Emotion-Aware Multimodal Systems for E-Learning Enhancement,” Frontiers in Artificial Intelligence, vol. 7, 2024.
[23] N. Hasan, “Context-Aware AI for Real-Time Student Support in Online Learning,” IEEE Transactions on Artificial Intelligence, vol. 5, no. 1, pp. 35–48, 2024.
[24] P. Nedungadi, G. Veena, K.-Y. Tang, R. R. K. Menon, and R. Raman, “AI techniques and applications for online social networks and media: Insights from BERTopic modeling,” IEEE Access, vol. 13, pp. 37389–37407, 2025.