

# EDU-SENSE: AI-Driven Adaptive E-Learning System for Primary Education

Nithiyasagar Shanmugalingam  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*  
Malabe, Sri Lanka.  
it21807794@my.sliit.lk

Roshni RJ  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*  
Malabe, Sri Lanka.  
it21831690@my.sliit.lk

Liseyanth Maheshwaran  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*  
Malabe, Sri Lanka.  
it21808920@my.sliit.lk

Uthpala Samarakoon  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*  
Malabe, Sri Lanka.  
uthpala.s@sliit.lk

Imana Rizvee  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*  
Malabe, Sri Lanka.  
it21237836@my.sliit.lk

Karthiga Rajendran  
*Faculty of Computing*  
*Sri Lanka Institute of Information Technology*  
Malabe, Sri Lanka.  
karthiga.r@sliit.lk

**Abstract**—This research addresses critical gaps in primary education by integrating AI-driven e-learning tools with traditional classroom teaching, ensuring technology supports rather than replaces teachers while improving student learning outcomes. In getting ready for high-pressure exams like Sri Lanka's Grade 5 Scholarship Examination, maintaining student engagement and emotional well-being remains a significant challenge. EDU-SENSE is an AI-driven web platform designed for primary education that integrates personalized content delivery, stress detection, and an emotion-aware chatbot. The system begins with a diagnostic pre-test to assess individual skill levels, and provides syllabus-aligned questions, and delivers targeted revision based on identified knowledge gaps. A stress-aware module and motivational chatbot provide emotional support, while peer collaboration features promote teamwork and knowledge sharing. Using Machine Learning, Natural Language Processing, and fine-tuned Large Language Models, EDU-SENSE offers a holistic solution that enhances both academic performance and emotional well-being in young learners.

**Keywords**—Large Language Models, Natural Language Processing, Stress detection, Peer collaboration, Machine learning.

## I. INTRODUCTION

The early years of education are critical to shaping the child's cognitive development, emotional intelligence, and academic foundation. During this stage, students develop essential thinking skills, learning habits, and perceptions of education. In Sri Lanka, the primary education phase (grades 1-5) is part of the 13-year national school system and serves as a crucial foundation for future academic success. However, this stage is often characterized by high academic pressure, particularly due to competitive examinations such as the Grade 5 Scholarship Examination [1]. Traditional teaching methods commonly adopt a "one-size-fits-all" approach, overlooking

differences in cognitive ability, emotional state, and learning preferences. This misalignment can lead to disengagement, increased stress, and reduced motivational factors that hinder meaningful learning and may contribute to long-term negative attitudes towards education.

Although e-learning platforms have grown in popularity, most focus on static content delivery with limited feedback loops and minimal emotional support. These systems often fail to adapt to changing learner's needs or facilitate the social interaction essential for early learners. In contrast, adaptive e-learning, powered by artificial intelligence (AI), can tailor educational content to individual learning styles, improving engagement, knowledge retention, and emotional connection [2], [3]. However, when implemented in isolation, such systems may lack the mentorship, personal touch, and collaboration found in traditional classrooms.

Integrating AI-driven adaptive e-learning with classroom-based instruction offers a promising hybrid model. AI enables personalized content adaptation, continuous performance monitoring, real-time stress detection, and targeted revision recommendations, while classrooms provide face-to-face guidance, emotional support, and peer collaboration. This combination can bridge learning gaps, boost engagement, and reduce the negative effects of academic pressure.

This research introduces 'EDU-SENSE', an AI-powered, emotionally aware, and socially engaging adaptive e-learning platform for primary school students. Rather than replacing traditional learning, EDU-SENSE complements traditional learning through intelligent digital tools that support a holistic approach [4]. The system delivers adaptive quizzes, monitors stress levels in real time, and, if a student is stressed, activates the chatbot, which aims to motivate the student, since motivation is considered a crucial factor in the academic

life of a student [5], [6]. The system integrates a content adaptation engine that tailors instructional materials to the individual skill level of each student, provides personalized revision recommendations, and promotes peer collaboration through an interactive chat environment. When integrated with traditional teaching, these capabilities create a balanced learning environment that promotes cognitive growth, emotional well-being, and sustained academic motivation.

*EDU-SENSE* integrates a content-adaptation engine that tailors instructional material to each learner's current skill, a performance-based revision recommender, and an interactive peer-collaboration space; when combined with classroom instruction, these capabilities promote balanced growth across cognition, emotional well-being, and sustained motivation. At its core, the platform couples an emotion-aware support layer using computer vision analysis of facial and behavioral cues to estimate stress and, when elevated, triggering an empathetic chat bot in text or voice to help learners regain focus, with an adaptive content engine driven by large language models that generates syllabus-aligned MCQ's to categorize students as highly, moderately, or weakly skilled and to update subsequent question sets accordingly. A performance-based recommended analyzes within-session patterns to pinpoint specific difficulties and proposes targeted, short revision activities, while the collaboration module forms groups by aligning interests and skill levels to encourage cooperative problem solving and peer support. Together, these elements create an inclusive, human-centered experience that personalizes content, responds to emotional states, and facilitates social learning and strengthening academic performance while nurturing resilience and curiosity; the remainder of this paper details the system design, architecture, and practical implications for scalable deployment among primary learners in Sri Lanka and similar contexts.

## II. LITERATURE REVIEW

Adaptive e-learning environments powered by Artificial Intelligence (AI) have shown notable improvements in student engagement and learning outcomes compared to traditional teaching methods. Systems that adapt content to individual learning styles report higher engagement across skills, performance, participation, and emotional involvement than static e-learning platforms [3]. Curriculum-aligned platforms expand access but seldom provide real-time personalization for primary learners. In Sri Lanka, the *Nenasa* portal offers syllabus-based lessons and past papers from Grade 1 upward, improving availability but providing limited adaptive or emotional support [7]. Similar "past-paper" applications, such as *Examiner*, automate MCQ practice and scoring yet lack fine-grained progress modeling and do not detect stress or deliver motivational feedback [8]. Adaptive learning systems personalize instruction by adjusting content to learner profiles and behavior. Trials with child-oriented math applications report significant gains over conventional methods, demonstrating the potential of tablet-based learning for early grades [9]. Game-based programs such as *NumberShire* also improve

foundational skills in primary education [10]. However, many platforms still rely on predefined content structures and cumulative summaries, limiting real-time personalization and topic-specific feedback [11], [12].

Adaptive systems employing item-response modeling improve the estimation of learner ability and question difficulty, enhancing diagnostic precision and drill efficiency [13]. Large-scale online curriculum evaluations similarly show positive outcomes even when adaptivity focuses mainly on pacing rather than continuous difficulty control [14]. Building on these insights, *EDU-SENSE* generates topic-specific quizzes in Easy, Medium, and Hard levels, continuously adjusting difficulty from learner responses to deliver timely, individualized support.

Affective computing strengthens intelligent tutoring by recognizing and responding to learner emotions. The *Affective AutoTutor* project demonstrated that combining emotion detection with motivational, speech-based feedback through tone, pitch, rhythm, and conversational modulation reduces frustration and enhances learning outcomes [15], [16]. Further studies confirm that motivational tone in spoken feedback increases persistence and performance, highlighting voice as an effective real-time encouragement tool [17]. Computer-vision models fine-tuned on children's facial expressions achieve strong accuracy on pediatric datasets [18], while interaction logs have been used to classify affective states during learning [19]. Few systems, however, combine real-time emotion detection with immediate voice-based feedback designed for adaptive learning in primary education. This gap motivates *EDU-SENSE*'s integration of facial stress detection with empathetic, voice-driven chatbot support.

Natural Language Understanding (NLU) is essential at the primary level, where learner inputs are often brief or ungrammatical. Recent studies show that instruction-tuned Large Language Models (LLMs) can interpret child responses, adjust explanations to age-appropriate readability, and generate scaffolded hints, thereby reducing teacher workload and improving accessibility [20], [21]. In assessment, LLMs trained with explicit scoring rubrics achieve near-human performance on criterion-based grading, though robustness and bias remain concerns [22]. Unlike systems that separate interpretation, readability control, and formative feedback, *EDU-SENSE* unifies these through an OpenAI fine-tuned LLM. The model autonomously interprets student answers, evaluates them against defined rubrics, and generates explanations suited to the learner's reading level. This removes the need for manual prompts students simply provide free-form responses, which the system grades and explains in real time. By abstracting prompt engineering, *EDU-SENSE* provides seamless, feedback-rich interaction tailored for young learners while leveraging the full capabilities of large language models.

Collaborative learning further enhances engagement and social development. Research shows that grouping learners by multidimensional profiles ability, behavior, and interests produces higher collaboration quality and group cohesion than random assignment [23], [24]. *EDU-SENSE* applies this

principle by forming three-member groups that mix high, medium, and low-skilled learners sharing similar interests, fostering peer support integrated with adaptive and emotion-aware learning. System scalability and reliability are equally critical. Stateless REST-based architectures minimize server dependencies and support low-latency interaction, aligning with the hardware constraints of primary classrooms [25]. For personalized revision, analytics methods such as knowledge tracing infer evolving mastery levels and guide next-step recommendations. In EDU-SENSE, post-quiz analytics detect topic-specific weaknesses and generate individualized revision plans, completing the cycle from adaptive questioning to mastery-focused reinforcement.

### III. METHODOLOGY

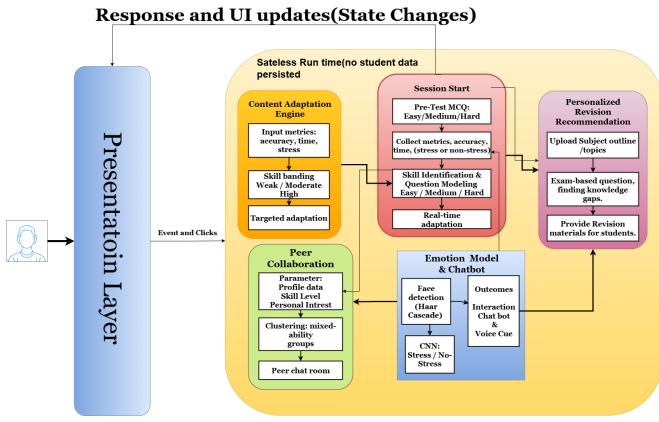


Fig. 1. Overall System Diagram and Architecture

EDU-SENSE is a web-based adaptive learning platform that integrates stress detection, intelligent assessment, personalized revision, and collaborative support for primary-level learners. The system uses a stateless architecture to protect privacy, with Python as the core backend and the OpenAI API powering a fine-tuned large language model (LLM). Students begin with a diagnostic pre-test, where responses are evaluated to classify them as weakly, moderately, or highly skilled. Based on this classification, EDU-SENSE adapts future content to match ability levels. Stress is continuously monitored through facial cues; when detected, a chatbot delivers motivational feedback through chat or voice cues to maintain engagement. During exam preparation, learners upload syllabus from which the system generates questions, evaluates responses, and recommends targeted revision resources. Finally, EDU-SENSE promotes collaboration by clustering learners into supportive peer groups. By uniting adaptive assessment, stress-aware support, revision guidance, and peer collaboration, EDU-SENSE offers a holistic, child-friendly environment.

#### A. Content Adaptation Engine

This component adapts the difficulty of learning content according to the learner's pre-test performance and classified skill level. The module ensures that learners receive content

aligned with their abilities while maintaining exam-like conditions.

*1) Pre-Test Phase:* At the beginning of a session, the learner selects their grade level, which determines the syllabus reference. A diagnostic pre-test with multiple-choice questions of varying difficulty is administered. Behavioral metrics recorded include:

- Accuracy ( $a_t \in \{0, 1\}$ )
- Response Time ( $\tau_t$  in seconds)
- Emotional State ( $e_t \in \{\text{stressed, not-stressed}\}$ )

*2) Skill Classification via K-Means:* Behavioral metrics are aggregated into normalized feature vectors:

$$x = [\bar{a}, \tilde{\tau}, \hat{e}]$$

where  $\bar{a}$  is mean accuracy,  $\tilde{\tau}$  latency, and  $\hat{e}$  encodes stress. A K-Means algorithm with  $k = 3$  partitions learners into Weak, Moderate, and High skill tiers, which determine the difficulty band for future content.

TABLE I  
PERFORMANCE-BASED CLASSIFICATION

Classification Category	Performance Threshold (% Correct)
Weakly Skilled	< 50%
Moderately Skilled	50% – 80%
Well-Skilled	> 80%

*3) AI-Driven Question Generation:* Based on the classified tier, exam-style questions are generated using a fine-tuned LLM. Prompts encode grade level, topic, and skill tier. Natural language processing techniques such as tokenization, lemmatization, and readability adjustment refine the generated items.

*4) Real-Time Adaptation:* A decision matrix maps accuracy ( $a_t$ ), response time ( $\tau_t$ ), and stress state ( $e_t$ ) to adjust the difficulty of the subsequent content.

#### B. Personalized Revision Recommendation

This module links each learner's weak areas to targeted revision material in a way that is simple to use and respectful of privacy. Learners upload a syllabus or textbook outline in PDF through the web interface; the Flask backend processes the file within the active session without permanent storage. PyPDF2 extracts the text, which is split into manageable units and normalized into topics using light LLM based labeling. For each topic the system generates short answer questions and evaluates responses with regular expression based keyword and phrase checks, lemmatization with weighted term scoring, and a GPT guided rubric when answers are ambiguous. Instead of returning a single score, the analysis surfaces specific gaps. For each gap the interface provides a direct reference to the relevant PDF section, a concise GPT generated summary, highlighted formulas or definitions, and a small set of focused practice questions. The learner's mastery profile is updated in real time within the session so the adaptive loop can immediately steer the next practice set, while the server retains no student data.

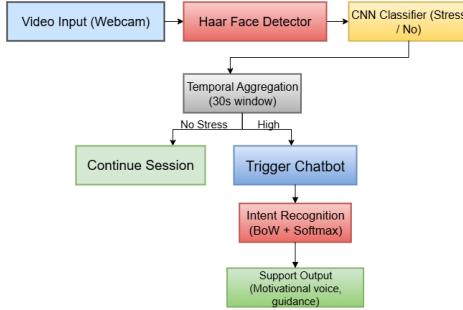


Fig. 2. Pipeline of the stress-aware chatbot system.

### C. Facial Stress Detection and Chatbot Support

1) *Facial Stress Detection*: This stage identifies when a learner appears stressed during a session. Webcam frames are captured in real time, and detected faces are analyzed by a Convolutional Neural Network (CNN) trained to classify *stressed* vs. *non-stressed* expressions. The model was trained on a dataset of 12,275 labeled facial images (9,795 for training and 2,480 for validation), resized to  $128 \times 128$  pixels and augmented through rotation, shifting, flipping, and brightness variation to enhance generalization.

The CNN architecture consists of three convolution–max pooling blocks (32, 64, and 128 filters), followed by batch normalization, a dense layer with 128 ReLU units and a dropout rate of 0.5, and a final sigmoid output neuron for binary classification. Training used the Adam optimizer with a learning rate of  $1 \times 10^{-3}$  and binary cross-entropy loss for 250 epochs, achieving approximately 80% validation accuracy.

To minimize false alarms, predictions are averaged over short time windows so that stress is flagged only when elevated probabilities persist across consecutive frames. The aggregated stress signal then triggers the **Chatbot Support Module**, which delivers brief motivational or calming prompts to help learners maintain focus and emotional balance.

To minimize false alarms, predictions are averaged over short time windows so stress is flagged only when high probabilities persist across consecutive frames. The aggregated stress signal triggers the Chatbot Support Module, which delivers brief motivational or calming prompts to maintain learner focus and emotional balance. **Evaluation**: The model was trained and validated using an 80–20 data split, with stratified sampling to maintain class balance. The training set consisted of 9,795 images, while 2,480 unseen samples were used for validation. Model performance was evaluated based on validation accuracy obtained during training and testing on the reserved dataset, achieving approximately 80% accuracy. This indicates reliable performance for real-time stress detection in classroom environments.

2) *Chatbot Support*: Once the system detects that a learner's stress is above a set threshold, the chatbot turns on. It offers support in text and in voice through the Web Speech API rather than giving corrections. Internally, an NLTK pipeline lowercases, tokenizes, and lemmatizes each message, which

is matched against a small intents JSON of tags, example phrases, and templates. Features are simple bag-of-words vectors; a compact Keras network with a softmax layer predicts the most likely intent and returns the template if its score passes a confidence limit. If not, a keyword-based fallback and a light context tracker steer the reply toward stress-support messages. Responses are short and age-appropriate, and no chat transcripts are stored. Together with the stress detector, this loop notices trouble quickly and offers calm encouragement so the learner can refocus. Together, the stress detection model and the chatbot form a closed loop: the former identifies moments of difficulty, while the latter steps in with immediate encouragement. This pairing ensures that learners are not left struggling in silence but are guided to re-engage with their studies in a positive and emotionally supportive way.

### D. Peer Collaboration through Clustering

This module facilitates collaborative learning by clustering students into balanced peer groups after individual skill classification. The objective is to enhance knowledge sharing and adaptive peer support by combining learners of varying proficiency levels.

1) *Data Preparation*: Learner profiles, including performance metrics and behavioral attributes, are preprocessed through feature scaling for numerical variables and one-hot encoding for categorical variables. A *ColumnTransformer* ensures standardized input representation.

2) *Cluster Formation*: To ensure effective collaboration, a K-Means clustering algorithm is employed. The optimal number of clusters,  $k$ , is determined via silhouette analysis. The silhouette score is computed as:

$$S = \frac{b - a}{\max(a, b)} \quad (1)$$

where  $a$  is the average intra-cluster distance and  $b$  is the nearest-cluster distance. Based on this metric, clusters are constructed such that each contains one well-skilled, one moderately skilled, and one weakly skilled student. This ensures diversity in ability levels within each group.

3) *Cluster Profiles and Visualization*: Principal Component Analysis (PCA) projects high-dimensional learner data into a 2D latent space for interpretability. Each cluster is summarized through descriptive profiles outlining study styles, error patterns, and engagement characteristics. Visualizations provide intuitive insights into the distribution of mixed-ability groups.

4) *Collaborative Chat Environment*: Each cluster is provisioned with a shared chat interface, enabling synchronous peer interaction. Within this environment, highly skilled learners can mentor weaker peers, while moderate learners serve as mediators, creating a structured knowledge exchange cycle.

5) *API Integration and Outputs*: A Flask-based API exposes endpoints for (a) assigning new learners into existing clusters, and (b) retrieving cluster profiles with both machine-readable (cluster ID) and human-readable descriptors. Operational outputs include cluster allocation, group-level profiles, and PCA-based visual representations. To preserve privacy, only anonymized embeddings and cluster identifiers are stored.

#### IV. RESULTS AND DISCUSSION

In this study, *EDU-SENSE* has been developed to motivate, assist primary learners, and help traditional teaching in Sri Lankan schools. The core aim is to deliver syllabus aligned questions drawn from standard educational resources, and reflect session level performance, and strengthen learners' enthusiasm for study.

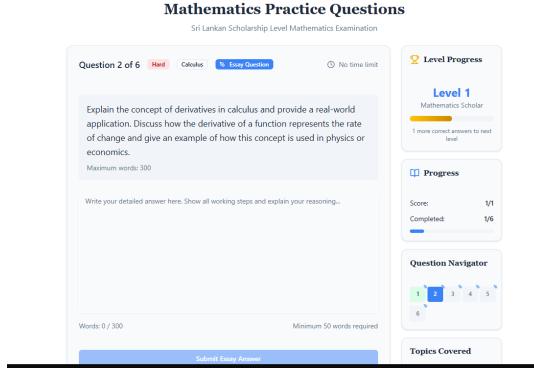


Fig. 3. UI of Question Answering

*Fig. 3* illustrates the answer workspace where students submit responses to the presented items. Consistent with the platform's stateless design, the system does not keep server side profiles; instead, a session level learner card shown in the browser can display grade and the current performance band, while a simple dashboard view summarizes marks from recent practice sets so students can review their progress during the session.

##### A. Emotion-Aware Chatbot

The emotion-aware chatbot was central to improving the learning experience by offering real-time motivational support. Unlike typical tutoring assistants, this chatbot does not provide direct academic instructions or revision tips. Instead, it monitors students' stress levels and delivers supportive responses in both text and voice formats. When students showed signs of stress, the chatbot generated calming, empathetic responses to reduce anxiety and help them refocus. During low-stress periods, it reinforced positive emotions with motivational feedback to maintain engagement. This balance of intervention and encouragement fostered a supportive learning environment where students felt emotionally connected rather than isolated. The use of both text and voice communication accommodated diverse learner preferences, enhancing inclusivity. The emotion-aware system's effectiveness relied on two key components the stress detection model and the chatbot interface working together to deliver responsive emotional support.

- 1) *Accuracy of the stress detection model:* The reliability of the chatbot depended strongly on the accuracy of the stress detection model. The system employed a Convolutional Neural Network (CNN) trained on a diverse dataset covering a wide demographic, from children to adults, labeled as "stressed" or "not stressed." This

diversity ensured robustness across facial variations, age groups, and environmental contexts.

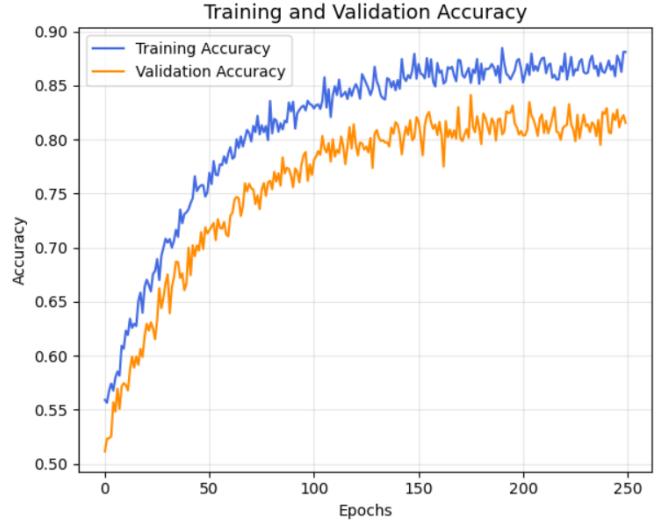


Fig. 4. Accuracy of the stress detection model (82%).

As shown in *Fig. 4*, the model achieved an accuracy of 82%, which is sufficiently high for practical classroom deployment. While occasional misclassifications occurred in low-light or highly dynamic settings, overall performance remained stable. The aggregation of predictions over 30-second windows further minimized noise and improved reliability of stress-level estimation.

- 2) *Chatbot Effectiveness:* The chatbot used stress predictions to adapt its motivational strategies in real time. Evaluations showed this approach reduced stress-related interruptions and extended students' engagement during long study sessions. Students found the chatbot to be a non-intrusive companion that made learning more balanced and enjoyable. The adaptive encouragement fostered persistence without overwhelming learners. The chatbot predictive model achieved an accuracy of 89%. While this result demonstrates promising effectiveness, the system is continuously being refined to improve both accuracy and precision. Future iterations are expected to further reduce misclassifications, ensuring more reliable support across diverse learning contexts.

##### B. Peer-Collaboration

The peer collaboration module was developed to encourage cooperative learning rather than isolated study. Groups were formed using a clustering model that considered skill level, age group, and personal interests, ensuring both academic balance and social compatibility. The students worked together on shared problem-solving tasks and short discussions. Stronger learners reinforced their understanding by explaining concepts, while weaker learners benefited from peer support in a low-pressure setting. Interest-based grouping also made discussions more natural and engaging. The evaluation results

showed that clustering with  $k = 3$  effectively represented the three skill categories used throughout the system. The resulting clusters demonstrated adequate separation and practical usability for real-time grouping. Students reported feeling more engaged and supported, with weaker learners showing increased participation during collaborative activities. The chat environment successfully enabled structured interaction within groups, while the clustering model performed reliably under the stateless architecture. Future work will focus on refining the clustering process by incorporating additional learner attributes and exploring persistent two-way communication to further enhance collaboration.

### C. Ethical Considerations and Data Handling

All testing was conducted with voluntary participation from 15 primary school students under parental supervision. No personally identifiable data or facial images were stored at any stage; all webcam frames were processed in real time and immediately discarded. The system design follows a stateless architecture to ensure privacy, and only aggregated, anonymized performance summaries were used for analysis. The study aligns with institutional ethical guidelines for child data protection and informed consent.

## V. CONCLUSION

The EDU-SENSE platform combines stress-awareness, adaptive content delivery, and an emotion-aware chatbot to support primary learners from Grades 1 to 5. It addresses everyday skill development as well as high-stakes exams like the Grade 5 Scholarship Examination in Sri Lanka. By adapting to students' cognitive abilities and emotional states, EDU-SENSE promotes engagement, motivation, and academic progress. Results show that integrating stress prediction with personalized content and motivational support boosts both confidence and learning outcomes. Future work will focus on expanding datasets and enhancing adaptive techniques to further optimize revision and exam preparation.

## ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Sri Lanka Institute of Information Technology (SLIIT) for providing continuous guidance, mentorship, and academic supervision throughout the research. The team also extends heartfelt thanks to Al-Hilal Central College and OKI International School for their valuable support and cooperation in allowing real-world testing and evaluation of the EDU-SENSE platform. Their collaboration and feedback were instrumental in validating the system's practical applicability within real classroom environments.

## REFERENCES

- [1] (2021) The Education System in Sri Lanka. Studycountry.com. [Online]. Available: <https://www.studycountry.com/guide/LKeducation.htm>
- [2] M. F. Odhaib, "Does e-learning give a better result than traditional learning," *International Journal of Computer Science and Mobile Computing*, vol. 7, no. 9, pp. 29–36, 2018.
- [3] H. El-Sabagh, "Adaptive e-learning environment based on learning styles and its impact on development students' engagement," *International Journal of Educational Technology in Higher Education*, vol. 18, 2021.
- [4] I. Bojovic and R. Antonijevic, "Students' motivation to learn in primary school," *Open Journal for Psychological Research*, vol. 1, no. 1, 2017.
- [5] F. Sherwani, M. Murtaza, J. Shamsi, Y. Ahmed, and M. Usman, "Ai-based personalized e-learning systems: Issues, challenges, and solutions," *IEEE Access*, vol. 10, pp. 81 323–81 342, 2022.
- [6] G. Wettergren, "Factors affecting student motivation in distance education," *Sub-theme: Quality and ODL*, 2011.
- [7] D. A. PLC, "Nenasa educational mobile app," Online; trilingual syllabus-aligned lessons, past papers, and quizzes for Grades 1–A/L, 2023, accessed: Aug. 10, 2025; Dialog Foundation partnership.
- [8] E. Lanka, "Sri lankan mobile application for mcq based past-paper practice," Mobile app (Google Play), accessed: Aug. 10, 2025.
- [9] L. A. Outhwaite, M. Faulder, A. Gulliford, and N. J. Pitchford, "Raising early achievement in math with interactive apps: A randomized control trial," *Frontiers in Psychology*, 2018.
- [10] I. of Education Sciences (IES), "Multi-site randomized controlled trial of numbershire level 1 gaming intervention," Award R324A160125, U.S. Department of Education; accessed Aug. 10, 2025, 2016.
- [11] A. Shemshack and J. M. Spector, "A systematic literature review of personalized learning terms," *Smart Learning Environments*, vol. 7, no. 1, p. 33, 2020.
- [12] E. du Plooy and S. Potgieter, "Personalized adaptive learning in higher education: A scoping review of key characteristics and impact," *Heliyon*, vol. 10, p. e39630, 2024.
- [13] S. Klinkenberg, M. Straatemeier, and H. L. J. van der Maas, "Computer-adaptive practice of maths ability using a new item response model for on-the-fly ability and difficulty estimation," *Computers Education*, vol. 57, no. 2, p. 1813–1824, 2011.
- [14] G. Abdelrahman, Q. Wang, and B. Nunes, "Knowledge tracing: A survey," *ACM Computing Surveys*, vol. 55, no. 11, pp. 1–37, 2023.
- [15] H. Wang and K. Woodworth, "A randomized controlled trial of two online mathematics curricula," *Society for Research on Educational Effectiveness*, 2011.
- [16] S. D'mello and A. Graesser, "Autotutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 2, no. 4, pp. 1–39, 2013.
- [17] S. D'Mello, N. Blanchard, R. Baker, J. Ocumpaugh, and K. Brawner, "I feel your pain: A selective review of affect-sensitive instructional strategies," *Design recommendations for intelligent tutoring systems*, vol. 2, pp. 35–48, 2014.
- [18] S. Wolff and A. Brechmann, "Carrot and stick 2.0: The benefits of natural and motivational prosody in computer-assisted learning," *Computers in Human Behavior*, vol. 43, pp. 76–84, 2015.
- [19] U. Laraib, A. Shaukat, R. A. Khan, Z. Mustansar, M. U. Akram, and U. Asgher, "Recognition of children's facial expressions using deep learned features," *Electronics*, vol. 12, no. 11, p. 2416, 2023.
- [20] M. H. Myers, "Automatic detection of a student's affective states for intelligent teaching systems," *Brain Sciences*, vol. 11, no. 3, p. 331, 2021.
- [21] H. Tran, Z. Yao, L. Li, and H. Yu, "Readctrl: Personalizing text generation with readability-controlled instruction learning," *arXiv preprint arXiv:2406.09205*, 2024.
- [22] F. Padovani, C. Marchesi, E. Pasqua, M. Galletti, and D. Nardi, "Automatic text simplification: A comparative study in italian for children with language disorders," in *Proceedings of the 13th Workshop on Natural Language Processing for Computer Assisted Language Learning*, 2024, pp. 176–186.
- [23] D.-W. Zhang, M. Boey, Y. Y. Tan, and A. H. S. Jia, "Evaluating large language models for criterion-based grading from agreement to consistency," *npj Science of Learning*, vol. 9, no. 1, p. 79, 2024.
- [24] G. Rolle-Greenidge and P. Walcott, "Group formation in a cross-classroom collaborative project-based learning environment," in *Proceedings of the 17th International Symposium on Open Collaboration*, 2021, pp. 1–6.
- [25] C. Liang, R. Majumdar, and H. Ogata, "Learning log-based automatic group formation: system design and classroom implementation study," *Research and Practice in Technology Enhanced Learning*, vol. 16, no. 1, p. 14, 2021.