Adaptive Online Learning Platform To Enhance Primary Education

Inuri Hapuarachchi
Department of Computer Science, Faculty of Data Science
Sri Lanka Institute of Information Technology
New Kandy RD, Malabe, Sri Lanka
inuri.chamathka@gmail.com

Panduka Wijesundara
Department of Computer Science, Faculty of Computing
Sri Lanka Institute of Information Technology
New Kandy RD, Malabe, Sri Lanka
panduka.s.wijesundara@gmail.com

Yushan Jayasekara
Department of Computer Science, Faculty of Computing
Sri Lanka Institute of Information Technology
New Kandy RD, Malabe, Sri Lanka
yushanjayasekara99@gmail.com

Dilshi Samarakkody
Department of Computer Science, Faculty of Computing
Sri Lanka Institute of Information Technology
New Kandy RD, Malabe, Sri Lanka
dilshi.samarakkody@gmail.com

Dharshana Kasthurirathna
Department of Computer Science, Faculty of Computing
Sri Lanka Institute of Information Technology
New Kandy RD, Malabe, Sri Lanka

Samadhi Rathnayake
Department of Computer Science, Faculty of Computing
Sri Lanka Institute of Information Technology
New Kandy RD, Malabe, Sri Lanka

***Abstract*— This research paper presents an innovative adaptive learning system tailored specifically for primary school students. Through a multi-faceted approach, this system integrates cutting-edge technologies and methodologies to address key challenges in primary education. The purpose of this study is to develop a comprehensive adaptive online learning platform that encompasses facial authentication, addressing emotional weaknesses, personalized recommendation systems, and real-time attentiveness monitoring. Firstly, a novel facial authentication system is introduced to streamline login processes and enhance the user experience by leveraging facial recognition technology, including Vision Transformers (ViT), Face Transformers, and established Convolutional Neural Networks (CNNs) like VGGFace2, FaceNet, and MobiFace. Secondly, the system incorporates a framework for classifying emotions in primary school children, augmented by a reinforcement learning (RL) model, and utilizing the cutting-edge transformer model Bidirectional Encoder Representations from Transformers (BERT). This framework aims to assist students in identifying true emotions and provides teachers with informed guidance on addressing emotional challenges. Thirdly, a recommendation system is developed to improve students' educational outcomes by administering subject-specific and general knowledge quizzes and offering tailored suggestions based on performance data and machine learning algorithms, such as knowledge graphs and transfer learning techniques. Lastly, the study proposes a unique approach to monitoring students' attentiveness in real-time during online quizzes using machine learning methods, including a pre-trained MobileNet model for transfer learning and real-time video processing algorithms. The significance of this research lies in its potential to revolutionize primary education by providing educators with powerful tools to support personalized learning and emotional growth. This adaptive learning system has far-reaching positive outcomes for the future of education, promoting inclusivity, engagement, and academic success for all primary school students.**

Keywords—Adaptive online learning system, Facial authentication, Facial recognition, Emotion classification, Recommendation system, Real-time attention monitoring,

I. Introduction

In today's educational landscape, the integration of technology has led to a new era of pedagogical methods, especially in primary education. As traditional classrooms give way to digital platforms, the imperative to ensure effective student engagement and address emotional needs becomes increasingly apparent. This introduction elucidates the multifaceted challenges that arise in online learning environments and proposes a comprehensive solution through the development of an adaptive online learning system tailored explicitly for primary school children.

The rapid expansion of online learning platforms has democratized access to education, offering personalized experiences and breaking down geographical barriers. However, transitioning to virtual classrooms poses challenges, especially in engaging young learners and supporting their emotional well-being. Primary school children face unique obstacles due to cognitive limitations and attention span constraints. One significant barrier is the conventional authentication method, which can be cumbersome for young learners. Addressing this, our research proposes a novel facial authentication system tailored for primary school children, aiming to streamline the login process and enhance the learning experience.

In addition to addressing authentication challenges, our research endeavors to tackle another critical aspect of primary school education: emotional well-being. Traditional educational approaches often overlook the nuanced emotional needs of young children, leading to unaddressed emotional challenges that may impede their learning journey. Recognizing the significance of emotional intelligence in academic success and overall development, our research aims to develop an adaptive learning system tailored explicitly for primary school students, encompassing an emotional weakness addressing component.

Several studies have explored affective recommender systems in virtual learning environments to enhance content recommendations and improve the learning process. Personalized recommendations tailored to individual student needs, expertise levels, and emotional states are essential for optimizing the learning experience. Our system addresses this need by providing a questionnaire-based on subject-wise general knowledge, enabling students to easily take quizzes and receive automated feedback on their educational performance. Leveraging a Large Language Model (LLM) and recommendation system, our approach guides students through personalized recommendations based on their educational levels, using Explainable AI technology to provide transparent reasoning behind each recommendation.

Moreover, recently online learning has gained prominence, necessitating solutions to maintain student engagement, and understanding in virtual classrooms. To address this, we advocate for integrating an attention monitoring system with an adaptive online learning platform. By utilizing transfer learning and machine learning models, our system analyzes video data to infer students' attention levels in real-time during online quizzes. This method provides educators with a unique tool to assess and address students' levels of involvement, ultimately enhancing the learning experience.

Creating an adaptive online learning platform tailored for primary school students could greatly enhance early childhood education. By tackling issues like login difficulties, attention spans, and emotional support, such a system can offer young learners a more interactive and nurturing learning space. Personalized learning opportunities may lead to better social and emotional development, ultimately fostering academic success. This study seeks to bridge a vital gap in educational technology by introducing an all-encompassing solution that utilizes advanced AI methods to cater to the specific requirements of primary school students, ultimately aiming to enrich learning experiences and boost educational achievements.

II. Literature Review

Our literature review delves into the landscape of adaptive online learning systems, particularly their application focusing on primary school students. Despite the growing interest, existing systems have received limited evaluations [13] for this age group, and accessibility remains a challenge. Our study aims to fill this gap by developing a tailored system, ensuring inclusivity, and addressing the unique needs of young learners.

Face recognition and authentication have been active areas of research in computer vision and artificial intelligence for decades, with early focus on hand-crafted feature extraction techniques and shallow machine learning models [1]. However, the advent of deep convolutional neural networks (CNNs) significantly advanced facial analysis, enabling more robust detection, alignment, and recognition [2], [3]. Recent years witnessed extensive research into enhancing facial recognition through attention mechanisms and novel architectures like Vision Transformers [5].CNNs have become the dominant approach, showcasing significant gains with architectures like DeepFace[2] and DeepID. Further advancements include the introduction of deeper networks like Visual Geometry Group (VGG) and Residual Network (ResNet), alongside innovative techniques such as triplet loss functions for learning embedding spaces [3]. Further techniques like center loss [4] and large margin softmax have also been proposed to enhance the discriminative power and generalizability of learned facial representations. Attention mechanisms[5] and Vision Transformers [6] augment CNN architectures to address limitations and enhance performance.

Despite advancements in CNNs, limitations persist in modeling long-range dependencies in faces, with attention mechanisms and Vision Transformers showing promise but requiring further research [5]. Challenges include high computational complexity and data requirements for pre-training ViTs, alongside the need for efficient ViT architectures tailored for facial recognition [7]. Additionally, achieving robust accuracy under unconstrained conditions and enabling explainable facial representations remain ongoing challenges. The methodology focuses on adapting attention mechanisms and Vision Transformers to address these limitations by exploring techniques for effective modeling of long-range dependencies and developing tailored ViT architectures [5], [6], [7]. Additionally, methods for enhancing accuracy under varied conditions and improving interpretability will be investigated, along with approaches to validate model fairness and mitigate demographic biases in facial recognition systems.

Emotion detection encompasses various approaches, including keyword-based methods utilizing lexicons like WordNet and the National Research Council (NRC) word-emotion lexicon identify specific emotion-related keywords within text and match them with predefined lists[8] , rule-based approaches establish logical and grammatical rules to detect emotions[8], machine learning-based such as Naive Bayes, Support Vector Machine (SVM), and decision trees methods categorize text into emotion classes[8], deep learning-based architectures like Bidirectional–Gated recurrent unit (Bi-GRU) and Long Short-Term Memory(LSTM) utilize neural networks for pattern recognition, and transfer learning-based methods facilitates emotion detection across diverse domains by reusing pre-trained models [8].

Despite extensive exploration, challenges persist in evaluating and applying emotion detection techniques for primary school students. Existing methods encounter difficulties with informal text [11], spelling mistakes, slang[10] [11] , and lack intensity information for emotions[11]. Additionally, handling multiple emotions within a sentence[10] and adapting pre-trained models to new domains[10] pose significant hurdles. Overcoming these limitations necessitates further research to develop robust emotion detection methodologies tailored explicitly for primary school students. Emphasizing usability, educational relevance and addressing these gaps is crucial for enhancing socio-emotional learning outcomes in primary education.

Leveraging RL algorithms allows for dynamically adapting advice to individual learning styles and preferences, fostering a more engaging and effective learning environment. Recent studies explore RL algorithms for fine-tuning Large Language Models (LLMs) in conditional text generation[12], interacting with dynamic black-box guide LLMs like GPT-3 to achieve superior performance compared to supervised learning and default RL algorithms like Proximal Policy Optimization (PPO)[12]. Our methodology involves developing a tailored emotion extraction model for primary school children, utilizing a BERT model augmented with reinforcement learning to accurately detect emotional weaknesses and generate personalized advices, aiming to enhance socio-emotional learning outcomes in primary education.

A number of studies have explored recommendation systems to enhance e-learning, often focusing on university students and leveraging their prior academic performance data [14], [15]. However, limited research examines adaptive systems tailored to improve educational outcomes specifically for primary students. This is an increasingly pertinent issue as artificial intelligence proliferates in early childhood education contexts [16]. Our research seeks to address this gap through an adaptive questionnaire and recommendation system catered to primary students. Administering subject-based quizzes, the system evaluates responses using large language models to gauge proficiency levels [17]. It then furnishes personalized materials aligned with students' aptitudes. Explainable AI clarifies the rationale behind each recommendation to illuminate the inner workings. This level of tailoring and transparency sets our method apart. Overall, it marks a valuable advancement in leveraging AI techniques to systematically enhance learning for younger students.

A common objective across recommender systems is maximizing accuracy [18]. However, our focus lies in effectively uplifting students needing improvement while also further enriching higher performers. The system promotes comprehensive educational growth. Technically, we integrate layered components spanning knowledge graphs, spaCy models, Neo4j databases and OpenAPI to enable robust analysis and queries [17]. The interplay of these elements supports granular insights into primary students' strengths and areas needing development. It then translates these insights into targeted materials for uplifted outcomes. While previous publications concentrated predominantly on e-learning broadly, our approach specializes in the unique needs of primary learners based on their cognitive constraints. This represents an emerging direction warranting greater exploration to fully unlock AI's potential in early childhood educational contexts.

Attention-aware systems are utilized in e-learning environments to monitor learners' attention levels in real time, aiming to enhance the learning experience through various approaches like face analysis and head posture characteristics, offering feedback to both students and instructors. These systems bridge the gap between traditional classroom settings and online learning platforms by utilizing techniques such as face detection, eye gaze detection, and facial expression analysis, although limited research exists on monitoring attention in offline settings with distant cameras capturing less defined facial features [19]. The suggested solutions, exemplified by the Students Attention Monitoring and Alert Model (S-AMAM) algorithm, strive for real-time, cost-effective monitoring accessible through web applications [20].

Implementing student attention monitoring systems in e-learning platforms is essential for enhancing the effectiveness of online learning by tracking attention in real-time [20], [21], [22]. These systems utilize facial recognition and expression analysis to combat decreasing focus levels among students, estimating attention using real-time webcam footage and combining physical and emotional measurements, but the training data set is limited [20], [21], [22]. Both verbal and non-verbal cues are employed to evaluate attention levels, with some studies focusing on non-verbal cues exclusively [20], [21], [22]. Transfer learning in deep learning involves using a pre-trained model to start a new task, like image classification [23]. By transferring knowledge, it improves performance and enables models to adapt to different datasets [23]. This approach reduces training time and enhances computational efficiency, making it popular for image classification[23].

III. METHODOLOGY

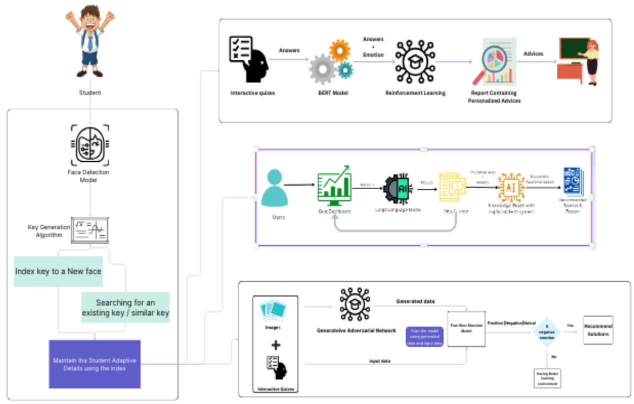

Fig. 1. Overall System Diagram of The Proposed Solution

The methodology section of this paper outlines a comprehensive approach to address the gaps identified in existing literature regarding adaptive online learning systems for primary students. This section delineates the systematic steps undertaken to design, implement, and evaluate our proposed solution, highlighting the integration of various technologies and methodologies to achieve the desired outcomes. Through a combination of quantitative analysis, algorithmic development, and qualitative evaluation, our methodology endeavors to provide actionable insights into improving the efficacy and accessibility of e-learning platforms for students, particularly those in primary education.

A. Using Vision Transformers for Facial Authentication

Face authentication is an important technology for user login and security. However, state-of-the-art face recognition models like Vision Transformers (VIT) can be computationally intensive, making them challenging to deploy on low-power devices like primary school computer systems. This paper proposes a novel face authentication approach using VIT models with optimized face extraction as a preprocessing step to improve processing speed.

a) Face Extraction:

The key idea is to preprocess input images to extract just the face region before feeding it into the VIT model. This avoids wasting computation on non-face regions. The steps are: Detects faces using a lightweight face detector like MTCNN. This gives bounding boxes around each face. Extract each face-bounding box and pad/resize to match the VIT input size. Feed only the extracted face regions into the VIT model. By removing background and other non-face areas, the VIT model only processes the relevant face area, improving computational efficiency.

b) VIT Face Authentication:

Preprocessed face images are fed into a VIT model pretrained on face recognition datasets like VGGFace2. Key hyperparameters and training procedures from state-of-the-art VIT face recognition research are utilized. Each user's face embedding is stored as their identity representation. Multiple images per user are enrolled to account for appearance variability. At test time, the input face image is embedded using the VIT model and compared to enrolled templates using cosine similarity. Thresholding distinguishes match vs non-match. The pipeline is evaluated on standard face recognition benchmarks. Processing time comparisons validate the improved efficiency of face extraction vs. feeding full frames into VIT. Recognition accuracy metrics ensure efficacy is maintained.

This work demonstrates an optimized face authentication methodology using VIT models. Face extraction as a preprocessing step reduces computational requirements, enabling deployment on primary school computer systems for user login.

B. *Addressing Emotional Weaknesses Using BERT and Reinforcement Learning*

a) Data Collection:

After consulting Dr. Indika Wijerathne, the Director of London NHS, a psychologist, and discussing with Mrs. Inoka Wijerathne, an English school teacher, we meticulously referred to the English syllabus for grades 4 and 5 to devise a set of questions aimed at gathering data for the emotion detection model. The questionnaire was structured into three sections: family, school, and friends & relatives. Each section comprised questions designed to elicit emotions such as happiness, sadness, feeling loved and cared for, fear, feeling unloved and uncared for. The survey was conducted with a cohort of 50 students encompassing both grade levels 4 and 5. Consequently, the final dataset comprises 50 records per emotion category.

b) Emotion Extraction Model:

We employed a methodology for emotion detection leveraging a BERT (Bidirectional Encoder Representations from Transformers) model. Initially, the dataset was preprocessed by encoding emotions into numerical labels and splitting it into training, testing, and validation sets. Utilizing the transformers library, the AutoTokenizer was employed to tokenize input texts, which were subsequently converted into tensors suitable for TensorFlow. The BERT model, obtained through TFBertModel, served as the backbone architecture, with additional layers added atop for emotion classification. Notably, GlobalMaxPooling1D and dense layers with ReLU and Sigmoid activations were integrated to facilitate feature aggregation and prediction. Following model compilation with Adam optimizer and categorical cross-entropy loss, training ensued, with training history recorded for analysis. Subsequent evaluation on the test set enabled mapping of predictions to original emotion labels. Additionally, a prediction pipeline was established to allow user input text for emotion prediction, alongside attention analysis to elucidate the model's decision-making process. This methodology stands out for its incorporation of BERT, contextualized word representations, and attention analysis, offering comprehensive insights into the emotion detection process.

Students were provided with three distinct quizzes focusing on family, school, and friends, designed collaboratively with the assistance of Dr. Wijerathne to accurately capture students' emotional states. These quizzes, each comprising 8-10 carefully crafted questions, allowed students the flexibility to complete them separately at their convenience. Notably, the quiz format accommodated both textual responses and the selection of pictures depicting various scenarios, ensuring a comprehensive assessment of emotional experiences. Following quiz administration, these sentences are fed into the BERT prediction pipeline to get the

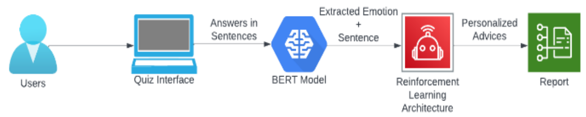

Fig. 2. Overall Structure of Emotional Weaknesses Addressing Model

true emotions and these extracted emotions together with the answers provided by the child was fed into the Reinforcement Learning Model as the input.

c) Reinforcement Learning Algorithm:

Our methodology involves training a Reinforcement Learning (RL) agent to generate personalized advice for handling emotions. The process begins by defining the state and action spaces, representing emotional states and possible actions, respectively. The RL agent employs the Q-learning algorithm, updating Q-values based on observed rewards and state transitions to maximize long-term rewards. To balance exploration and exploitation, an ε-greedy exploration strategy is implemented in the policy function. During training iterations, the agent interacts with the environment, selecting actions, generating advice prompts, utilizing a language model to generate advice, collecting feedback, and updating Q-values based on feedback. Feedback from users determines the rewards received by the agent, guiding its learning process. The advice generation process focuses on creating emotionally relevant prompts within a character limit. Our implementation utilizes Python and NumPy, with a language model for advice generation. Evaluation metrics include average reward and convergence, assessing the agent's effectiveness in providing helpful advice in diverse emotional scenarios. In summary, our methodology offers a systematic approach to training an RL agent for personalized advice generation, aiming to provide valuable support to individuals facing emotional challenges.

C. *Recommendation system to improve educational performance*

a) Questionnaire:

This system is designed to revolutionize the way students learn and excel academically. The online questionnaire system uses a large language model (LLM) accessed through the API key. This system assesses students' real-time educational levels and dynamically selects quizzes from relevant subjects based on their preferences. Using LLM, evaluating students' quiz responses, this system provides immediate feedback on their performance with educational level and results. Then pass that educational level to the knowledge graph to get recommendations to get better learning outcomes and improve students' educational levels.

b) Knowledge Graph for recommendation:

At the heart of this platform is a powerful knowledge graph that captures a wide array of educational topics and

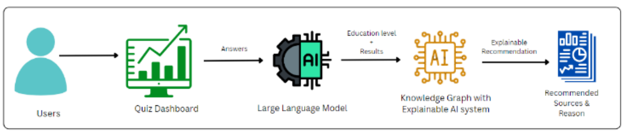

Fig. 3. Overall Structure of Recommendation System

concepts. We utilized the Neo4j database to construct this knowledge graph and the Resource-Action-Goal (RAG) network model to capture the hierarchical relationships between different educational domains. This enables us to recommend relevant study materials and resources to students, making learning more efficient and effective.

To ensure that our knowledge graph stays up-to-date and relevant to students' educational needs, we integrated spaCy-llm, a large language processing tool that extracts entities and relationships from the educational content. This tool automatically enriches our knowledge graph, making it a constantly evolving resource that adapts to students' needs. In addition, this system uses an explainable AI to take an explanation of the recommendations.

D. *Monitoring Sudents' Attentiveness*

a) Data Collection:

While collecting data, students are engaged in an online quiz, and while they are taking the quiz, a webcam captures student behavior and collects a video dataset from each student.

b) Preprocessing:

- OpenCV enables video capture via a webcam.
- Images were extracted from a recorded video stream at one frame per second.

Each picture is resized to 40x40 pixels and standardized (divided by 255).

c) Model Development:

Use the MobileNet pre-trained model to perform transfer learning. MobileNet is suitable for both mobile and embedded vision applications. To maintain learned features and avoid overfitting, freeze the majority of the MobileNet model's layers. Retrain the final 5 layers based on our classification job, using a trial and error strategy. Use approaches such as data augmentation to improve generalization.

d) Real-Time Processing:

When the quiz begins, the system will immediately take footage from the camera. Use a face recognition algorithm to extract facial information from each frame of the video. Extract and preprocess the facial area from each frame. Feed the preprocessed face photos into the trained MobileNet model, which will predict attention levels. Aggregate predictions across time to provide a more consistent portrayal of attention levels throughout the quiz.

IV. Results and Discussion

Advices generated through LLM after training the Reinforcement Learning agent on the sentences and their emotion.

Sentence	**Emotion**	**Action chosen by agent**	**Advice Generated by LLM**
I feel sad when my parents forget about important school events like parent-teacher conferences.	Sad	comfort	Acknowledge their feelings, reassure them of their worth, and involve them in activities that boost their confidence.
Mom listens to my dreams and helps me plan how to make them come true.	Loved & Cared	none	Foster a space where dreams are shared and planning for them is a collective effort.
I feel sad when my friends make plans without inviting me, making me feel like I'm not part of the group.	Sad	comfort	Acknowledge their feelings, "It's okay to feel sad. Let's talk about it. Remember, your worth isn't defined by others' actions.
Taught me about society.	Happy	advise	Use positive reinforcement to foster a love for learning and understanding the world around them.

V. Limitations and Challenges

While this research offers valuable innovations in adaptive e-learning for primary students, certain limitations remain to be addressed:

1) Facial Authentication:

Despite recent progress, facial recognition models still confront challenges including computational complexity, substantial data demands, overfitting, interpretability issues, and vulnerability to attacks. Moreover, real-world latency and resource constraints limit viable model complexity. Significant research remains to enhance efficiency, robustness, and fairness while reducing data needs and bias.

2) Emotional Weakness Addressing:

The emotional weaknesses addressing model is constrained by several limitations in its data collection process. Firstly, data were gathered from a small number of students, potentially limiting the generalizability of findings. Despite efforts to streamline questionnaires to 10 questions, this brevity may not fully capture the complexity of emotional states. Moreover, data were collected exclusively from one school, potentially introducing biases related to its specific demographics and environment. These limitations emphasize the need for caution when interpreting results and suggest avenues for future research to enhance the model's robustness.

3) Recommendation System:

Quality educational data availability poses a challenge in training high-performing machine learning models. Moreover, learning levels depend on various factors like styles and backgrounds, not solely quiz performance. Highly interpretable models often sacrifice predictive accuracy. Advanced explainability methods are warranted so recommendations resonate at the individual level. Additionally, large-scale knowledge graph implementation barriers persist.

4) Attention Monitoring system:

The investigation into real-time attention monitoring in education has several limitations. The size of the sample of student films employed may not accurately reflect the wide-ranging student population, thereby restricting the generalizability of the findings. The human annotation of ground truth data is subjective and has the potential to differ depending on the observer. The intricacy and expense of AI technologies for data protection can present challenges, as not all businesses can effectively adopt and utilize these systems.

In summary, facial authentication, emotion classification, and recommendation systems at a primary education level continue to face technological and data-related obstacles. Further research into efficient architectures, diverse data, and explainability is critical to unlocking their full potential while minimizing bias, resources and ethical risks. A concerted effort by cross-disciplinary teams could help overcome these limitations through responsible innovation.

VI. Future Work

A. *Facial Authentication*

Continuing research is warranted into adaptive facial recognition models incorporating spatial, channel and self-attention mechanisms to focus on informative regions while remaining lightweight. Validating fairness and minimizing demographic biases is also an imperative as adoption spreads.

B. *Emotional Weakness Addressing*

Collecting diverse emotional data from multiple schools across urban and rural areas can better capture the spectrum of students' backgrounds to enhance model flexibility. Advanced reinforcement learning techniques like multi-agent systems also offer promise.

C. *Recommendation System*

Further personalization of knowledge graph-based recommendations is viable, potentially aided by AI chatbots for customized student queries. Reinforcement learning presents another avenue for an increasingly tailored system.

D. *Attention Monitoring System*

Future research should explore variables like student posture and gamification's effect on attentiveness, assess the effectiveness of real-time attention monitoring systems in various educational settings, and investigate their impact on student engagement, learning outcomes, and overall educational experience.

Overall, refinements in facial authentication efficiency/fairness, emotional breadth via diverse data, and cutting-edge reinforcement learning personalization in recommendations remain open challenges. While this research put forth an innovative adaptive platform, augmentations in these domains could overcome limitations and better promote comprehensive development. As online learning permeates primary education, developing robust and ethical AI systems to elevate young minds remains an inherent responsibility. This research marks significant progress, setting the stage for impactful multi-disciplinary collaborations between technology creators, emotional intelligence experts, educators and policymakers to fully actualize AI's potential in this sphere.

Acknowledgment

We sincerely appreciate the guidance we received from everyone during this research project. We would like to express our gratitude to Mr. Samadhi Rathnayake, our supervisor and Dr. Dharshana Kasthurirathna, our co-supervisor, for their invaluable advices, feedback, and support throughout the entire project. Their guidance and thoughts were essential in enabling us move forward with this project.

Furthermore, we express our gratitude to Dr. Lakmini Abeywardhana for her insightful suggestions made during research presentations, which led us to consider crucial aspects leading to significant enhancements. We also thank Mrs. Inoka Wijerathne, the English teacher, and Dr. Indika Wijerathne, the psychologist, for lending their specialist knowledge to the development of the system's emotional weakness addressing component. Their perspectives as professionals in these disciplines added greater depth.

The contributions of all these individuals were instrumental in shaping and improving this research. We could not have successfully executed this extensive project without their selfless support. We will remain grateful for their guidance through every milestone in our academic journey.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld, "Face recognition: A literature survey," ACM Computing Surveys, vol. 35, no. 4, pp. 399-458, 2003.

[2] Y. Taigman, M. Yang, M. Ranzato and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701-1708.

[3] F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 815-823.

[4] Y. Wen, K. Zhang, Z. Li and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in Proceedings of European Conference on Computer Vision, 2016, pp. 499-515.

[5] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang and X. Tang, "Residual attention network for image classification," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 6450-6458.

[6] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold and S. Gelly, "An image is worth 16x16 words: Transformers for image recognition at scale," arXiv preprint arXiv:2010.11929, 2020.

[7] R. Meher, A. Singh, A. Nath and A. Prakash, "Selective face transformer," in Proceedings of International Joint Conference on Neural Networks, 2021, pp. 1-8.

[8] S. Kusal, S. Patil, K. Kotecha, R. Aluvalu, and V. Vijayakumar, "AI Based Emotion Detection for Textual Big Data: Techniques and Contribution," *Big data and cognitive computing*, Sep. 09, 2021. https://www.mdpi.com/2504-2289/5/3/43

[9] F. Ghanbari-Adivi and M. Mosleh, "Text emotion detection in social networks using a novel ensemble classifier based on Parzen Tree Estimator (TPE)," *Neural Computing and Applications*, vol. 31, no. 12, pp. 8971–8983, May 2019, doi: 10.1007/s00521-019-04230-9.

[10] A. Balahur and M. Turchi, "Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis," *Computer Speech & Language*, vol. 28, no. 1, pp. 56–75, Jan. 2014, doi: 10.1016/j.csl.2013.03.004.

[11] E. Batbaatar, M. Li, and K. H. Ryu, "Semantic-Emotion Neural Network for Emotion Recognition From Text," *IEEE Access*, vol. 7, pp. 111866–111878, 2019, doi: 10.1109/access.2019.2934529.

[12] W. Chen, J. Tian, C. Fan, Y. Li, H. He, and Y. Jin, "Preference-Controlled Multi-Objective Reinforcement Learning for Conditional Text Generation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 11, pp. 12662–12672, Jun. 2023, doi: 10.1609/aaai.v37i11.26490.

[13] B. Pardamean, T. Suparyanto, T. W. Cenggoro, D. Sudigyo, and A. Anugrahana, "AI-Based Learning Style Prediction in Online Learning for Primary Education," *IEEE Access*, vol. 10, pp. 35725–35735, 2022, doi: 10.1109/access.2022.3160177.

[14] P Ramya and S G Balakrishnan and M Kannan, *Recommendation system to improve students performance using machine learning,* Tamil Nadu, India: Department of Computer Science and Engineering, Mahendra Engineering College, 2020.

[15] Qian Zhang, Jie Lu, Guangquan Zhang, *Recommender Systems in E-learning,* Sydney: University of Technology, 2020.

[16] Jiahong Su and Weipeng Yang, *Artificial intelligence in early childhood education: A scoping review*, China: Faculty of Education, The University of Hong Kong, 2021.

[17] Exploring Knowledge Extraction Techniques for System Dynamics Modelling: Comparative Analysis and Considerations

[18] Fergyanto Efendy Gunawan, *A Learning Recommendation to Improve Electronic Textbook Learning Experience,* Vol.13. Japan: ICIC EXPRESS LETTERS, 2019.

[19] X. a. Y. J. a. L. J. a. Z. H. a. S. H. Ling, "A deep-learning based method for analysis of students' attention in offline class," Electronics, vol. 11, p. 2663, 2022.

[20] M. L. a. M. D. a. J. J. D. Roy, "Students attention monitoring and alert system for online classes using face landmarks," in 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), 2021, pp. 1--6.

[21] R. a. S. S. a. P. V. a. K. S. RK, "Real-time Attention Span Tracking in Online Education," in 2020 IEEE MIT Undergraduate Research Technology Conference (URTC), 2020, pp. 1-4.

[22] Z. a. A. F. a. P. M. M. A. a. G. M. a. A. L. Trabelsi, "Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student's Behavior Recognition," Big Data and Cognitive Computing, vol. 7, p. 48, 2023.

[23] J. H. a. D. R. a. T. S. D. a. J. H. a. N. N. a. M. A. a. N. S. Dewan, "Image Classification by Transfer Learning using Pre-Trained CNN Models," 2023.