Speech Processing with Integration of Deep Learning for Sustainable Development: An Overview

[1]Rupali Khaklary, Research Scholar
Computer Science, Krishna Kanta Handiqui State Open University, Guwahati-781017, Assam

***Abstract* – This paper explores the integration of Deep Learning with speech processing to enhance speech recognition and extract complex features from speech data. By leveraging advanced machine learning techniques, the study addresses the Sustainable Development Goal of providing quality education through energy-efficient and eco-friendly software practices. This innovative approach not only improves the accuracy and efficiency of speech recognition systems but also contributes to the global agenda of sustainable development by minimizing environmental impact. The findings of this study demonstrate significant improvements in speech recognition accuracy and processing speed, which can transform educational technologies and promote lifelong learning. The research underscores the potential impact of these advancements on sustainable development and quality education, providing insights into the challenges and future directions of developing such systems.**
Keywords: **Deep learning, Sustainable development, speech processing, sustained learning, lifelong learning, quality education.**

I. INTRODUCTION

In the modern world, sustainable development has become a paramount concern, encompassing environmental protection, social equity, and economic growth (UN, 2020). Achieving a balance between these elements requires innovative solutions and the integration of advanced technologies. One such technological advancement is deep learning (DL), which has shown significant promise in various fields, including speech processing [1,2]. Speech processing, which involves the analysis and synthesis of speech signals, plays a crucial role in numerous applications such as speech recognition, text-to-speech synthesis, and emotion recognition [3]. Traditionally, these tasks relied on handcrafted features and rule-based systems, which were limited in their ability to capture the complexity of human speech [4]. However, with the advent of DL, it has become possible to automatically extract intricate features from speech data, leading to substantial improvements in performance and accuracy [5]. Despite these advancements, the integration of DL in speech processing for sustainable development remains underexplored. Sustainable computing, which focuses on efficient and eco-friendly use of computing resources, is critical in this context [6,7]. As digital society evolves, the demand for energy-efficient and resource-conscious technologies grows [8]. The challenges lie in developing speech processing systems that not only perform well but also adhere to principles of sustainability [9]. Moreover, learning and knowledge acquisition are essential in human life

which help improve health outcomes, economic growth, gender equality, and reduced inequalities that are integral to sustainable development. Quality education, one of the Sustainable Development Goals (SDGs), underscores the importance of learning as a lifelong process. The integration of DL with speech processing offers promising avenues for advancing educational technologies, making learning more accessible, personalized, and engaging. Speech technologies can significantly enhance these processes by providing intuitive and accessible interfaces for information retrieval and analysis [11, 12]. The primary problem addressed in this paper is the need for integrating DL with speech processing to support lifelong and sustained learning (SDG 4). SDG 4 aims to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all [13]. Lifelong learning refers to the continuous pursuit of knowledge and skills throughout an individual's life, while sustained learning emphasizes consistent and ongoing educational engagement. This paper explores the current state of DL in speech processing, identifies the challenges and opportunities in this field, and discusses the potential for these technologies to support the learning process for sustainable development. By examining recent advancements and ongoing challenges, this study aims to highlight the critical role that speech processing integrated with DL can play in fostering lifelong and sustained learning to achieve equity and inclusive education. The goal is to provide a comprehensive overview of the field and propose directions for future research and development that align with sustainability principles [2].

II. Background and Motivation

A. Overview of Sustainable Development Goals

The Sustainable Development Goals (SDGs), adopted by all United Nations Member States in 2015 as part of the 2030 Agenda for Sustainable Development, represent a holistic approach to balancing human development and environmental stewardship. Among the SDGs, education (Quality education) is the central theme. The fourth goal of SDGs aims to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all". This goal highlights the pivotal role education plays in sustainable development, empowering people with the knowledge, skills, and values necessary to shape a more sustainable future. Education and lifelong learning are pivotal in achieving SDGs to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.

B. Importance of Quality Education and lifelong-learning

Quality education is fundamental to achieving many other SDGs. It is linked to improved health outcomes, economic growth, gender equality, and reduced inequalities. Education fosters critical thinking, problem-solving, and the ability to adapt to changing circumstances, all of which are essential for addressing the complexities of sustainable development. Lifelong and sustained learning encompasses learning opportunities across all ages and

Contexts, from early childhood education to adult learning and professional development with consistent and ongoing educational engagement for the future. It ensures that individuals can continuously upgrade their skills and knowledge, keeping pace with the rapid changes in technology and the global economy [14].

C. Importance of Sustainable Computing in achieving Lifelong Learning and Quality Education

Sustainable computing is crucial for achieving lifelong learning and inclusive education, key components of SDG 4. It focuses on the efficient and eco-friendly use of computing resources, which is critical for reducing energy consumption and minimizing environmental impact [8]. Integrating speech processing and DL within this framework transforms education systems promoting equitable and widespread distribution of resources.

Energy Efficiency and Resource Optimization: Sustainable computing ensures educational tools powered by speech processing and DL are sustainable and scalable [15] [16]. Energy-efficient DL models support adaptive learning platforms providing personalized learning experiences continuous skill development and knowledge acquisition [17,18,19,20,21].

Scalable Educational Solutions: Sustainable computing enables the development of scalable educational solutions that can reach a large number of learners, regardless of their geographic location or socioeconomic status [5]. By minimizing resource demands, sustainable computing enables scalable educational solutions, making them affordable and accessible to learners everywhere that support lifelong learning.

Inclusive Education Through Speech Processing with Sustainable Computing

- *Accessibility and Inclusivity:* Speech processing with DL and sustainable computing is crucial for creating inclusive educational environments. Speech recognition and text-to-speech (TTS) systems ensure that students with visual or hearing impairments can access educational materials and participate fully in learning activities [22,23].
- *Multilingual Support:* Sustainable computing and advanced speech processing enable multilingual systems that support diverse linguistic communities. These systems can provide educational content in multiple languages, fostering inclusive education while maintaining energy efficiency [24].

Equity in Education: The integration of speech processing and DL within a sustainable computing framework makes high-quality educational resources accessible to the underserved. Remote learning platforms powered by these technologies ensure equitable education opportunities for students in remote areas, essential for achieving SDG 4 [25].

Quality Education Through Sustainable Computing

- *Enhancing Educational Quality:* Sustainable computing improves education by developing advanced, environmentally responsible educational tools. DL models can analyze huge speech data to provide insights into learning patterns, helping educators tailor strategies to individual needs, and ensuring personalized support and feedback [11].
- *Innovative Teaching Methods:* Sustainable computing fosters innovative teaching methods, such as interactive voice-controlled tools and virtual classrooms powered by speech-processing technologies [26]. These methods enhance learner engagement, active learning, and critical thinking.
- *Proactive Interventions:* Speech-driven analytics supported by sustainable computing, enable early detection of learning difficulties or disengagement. By analyzing speech patterns and interactions, educators can proactively provide the necessary support, ensuring all learners have the opportunity to succeed [26].

D. Unique Role of Speech Technologies in Achieving Sustainability

Speech technologies, powered by deep learning, play a unique role in supporting sustainable development by enhancing communication, education, and decision-making processes [11]. These technologies offer intuitive and accessible interfaces for various applications, contributing to sustainability goals

III. Novel Applications of Speech Processing to Transformative Impact on Quality Education

Speech technologies can transform education by accessible learning tools and enhancing the quality of education. These technologies support lifelong, which is a key aspect of SDGs. The integration of DL in speech processing has enabled the development of advanced educational tools that cater to diverse learning needs. This integration has a transformative impact on education, enhancing the quality and accessibility of learning resources supporting lifelong learning, and contributing to SDGs by fostering an inclusive and equitable education system. For example, speech recognition technologies can facilitate language learning by providing instant feedback on pronunciation and fluency [27]. Deep learning-powered Text-to-speech (TTS) systems, can transform written texts into spoken words, thereby making educational resources more accessible to learners with visual impairments or reading challenges [28,29]. Moreover, DL-enhanced speech technologies can support personalized learning experiences. Adaptive learning platforms utilize speech processing to understand and respond to individual student needs, providing customized feedback and recommendations [30]. This personalization fosters an inclusive learning environment, addressing the unique challenges faced by each learner.

IV. Role and Potential of Speech Processing with Deep Learning in Lifelong and Sustained Learning

Enhanced Accessibility: Speech processing technologies, such as automatic speech recognition (ASR) and Text-to-Speech (TTS), significantly enhance accessibility in lifelong learning. These technologies allow individuals with visual or auditory impairments to access educational content, participate in discussions, and complete assignments through voice commands and auditory feedback.

Speech Recognition and Automated Assessment: Speech-processing applications are instrumental in language learning. Speech recognition technologies, powered by DL, facilitate language learning by providing instant feedback on pronunciation and fluency [27]. These tools can recognize and correct speech patterns, helping learners improve their language skills efficiently and engage in interactive conversations with AI-powered language tutors. These tools make language acquisition more engaging and effective, supporting lifelong learning objectives [24]. Automated speech assessment tools can evaluate oral exams, presentations, and language proficiency tests, reducing the burden on educators and ensuring consistent and objective grading [31,32,33,34]. In an automatic assessment of L2 spoken language proficiency, input sequential data from a learner is used to predict a grade or a level concerning holistic proficiency or specific aspects of proficiency [35,36]. This automation allows for scalable assessment methods, making quality education more accessible.

Text-to-Speech (TTS) Systems (Accessibility & Multilingual Support): TTS systems convert texts into speech, providing access to educational materials for visually impaired learners and those with reading difficulties. They also support multiple languages, aiding non-native speakers and promoting inclusivity and linguistic diversity in education [24].

Personalized Learning (Adaptive Learning Platforms & Real-Time Integration: DL-enhanced speech technologies enable adaptive learning platforms to understand and respond to individual student needs. These platforms can analyze speech data to tailor feedback and recommendations, providing a personalized learning experience [19]. This personalization addresses diverse learning styles and paces, ensuring that each student can achieve their full potential. Interactive voice response systems, powered by DL, allow for real-time interaction between learners and educational platforms. These systems can answer questions, provide explanations, and guide learners through complex topics, creating an engaging and dynamic learning environment [26].

Remote and Distance Learning (Virtual Classrooms & Interactive Educational Content): Speech-processing technologies facilitate remote learning by enabling voice-activated virtual classrooms and education platforms by enabling seamless communication between educators and learners. AI-powered speech recognition and synthesis tools ensure clear and accurate communication, even in noisy environments or with low-quality audio inputs [5]. These systems provide flexible and convenient learning opportunities for individuals who may not have access to traditional educational institutions due to geographic, financial, or time constraints [11]. This capability is crucial for maintaining the quality of education in remote and distance learning settings. Interactive voice-controlled educational content can enhance remote learning

experiences. Students can interact with digital textbooks, educational games, and simulation tools using natural language, making learning more engaging and effective. This positive learning experience cultivates sustained motivation and an active approach to learning [37].

Speech-Driven Analytics (Educational Insights & Early Intervention): Analysing speech data from classrooms and educational interactions can provide valuable insights into student engagement, participation, and comprehension [11]. Educators can use these insights to tailor their teaching strategies and improve the overall learning experience. Speech analytics can identify early signs of learning difficulties or disengagement, allowing educators to intervene promptly and provide the necessary support [26]. This approach helps ensure that all students receive the assistance they need to succeed.

V. Innovative Integration of Deep Learning in Lifelong and Sustained Learning

The integration of DL in speech processing is transforming education by promoting lifelong and sustained learning. This synergy enhances speech technologies while promoting sustainable practice [2]. DL enables the development of energy-efficient, resource-conscious systems that outperform traditional methods reliant on handcrafted features and ruled-based systems [4],[5]. By automatically extracting complex features, DL improves performance and offers personalized, accessible, and engaging educational experiences. This section examines innovative uses of these technologies to support continuous learning.

Personalized Learning Experiences: Personalization is essential for lifelong and sustained learning. DL algorithms analyze learners' speech data to understand individual learning patterns and needs, allowing educational systems to adapt content accordingly. Intelligent tutoring systems use DL to provide customized feedback, adjust difficulty levels, and suggest resources, creating engaging and effective learning experiences. By analyzing speech patterns, DL models identify areas needing support or challenge, optimizing the learning journey [2],[38].

Real-time Assessment and Feedback: DL models can perform real-time assessments of learners' performance. By analyzing speech patterns, responses, and interactions, these models provide instant feedback, helping learners identify areas for improvement and track their progress over time [26].

Interactive Learning Environments: DL enables interactive learning environments where learners can engage with AI tutors, participate in simulations, and explore complex concepts through voice interactions. The use of mobile devices, smartboards, MOOCs, tablets, laptops, and virtual labs has transformed education [39]. Integrating IoT in educational tools offers a cost-effective approach to high-quality learning [40,41,42]. These innovations promote active learning and critical thinking, essential for lifelong learning.

Speech-Activated Learning Tools: Speech-activated learning tools, powered by DL, provide intuitive, hands-free interaction with educational content, aiding learners with disabilities. Voice-activated assistants help navigate platforms and perform tasks via voice commands, promoting an inclusive learning environment [11].

Language Learning and Pronunciation Training: DL in speech processing significantly enhances language learning by accurately assessing pronunciation and fluency, and providing detailed feedback [24]. These systems simulate natural conversations for practice aiding proficiency [24]. Additionally, DL models support multiple languages, facilitating bilingualism and new language acquisition.

Emotion Recognition for Enhanced Learning Support: Recognizing learners' emotions through speech using DL models can enhance educational interactions by adapting strategies based on students' engagement and well-being. If frustration or disengagement is detected, the system can adjust to re-engage and support the learner, fostering a positive learning environment [2].

Interactive and Immersive Learning Environments: Integrating DL in speech processing enhances learning environments through Virtual reality (VR) and augmented reality (AR) platforms. These technologies, utilizing speech recognition and synthesis, enable learners to interact with virtual objects and scenarios using natural language, making learning, more engaging and effective, especially for complex subjects [43],[44].

Support for Remote and Blended Learning: DL-enhanced speech-processing technologies are essential for remote and blended learning. Real-time speech recognition transcribes lectures and discussions, making them accessible to remote learners and facilitating virtual communication. This integration ensures equitable access to learning opportunities for all students [11].

VI. SUSTAINABLE PRACTICES IN EDUCATION

Energy-Efficient Models: Developing energy-efficient DL models for speech processing minimizes the environmental impact of deploying these technologies on a large scale [15]. Techniques such as model pruning and quantization help reduce the computational resources required, making speech technologies more sustainable

Edge Computing: Implementing edge computing in educational technologies allows for local data processing, reducing the need for extensive data transfer and thereby decreasing energy use [16]. This approach ensures that speech processing applications are both efficient and environmentally friendly.

Open Educational Resources (OER): Speech processing technologies can enhance OER by providing high-quality, accessible educational content to a global audience. This democratization of knowledge supports lifelong learning and contributes to sustainable education practices.

Scalable Solutions: Scalable speech processing solutions ensure that educational tools can reach a wide audience without compromising on quality or performance [5]. This scalability is essential for addressing the educational needs of diverse populations, particularly in under-resourced regions.

VII. Recent Advancements in Speech Processing and Deep Learning

Improved Speech Recognition: Advances in DL, particularly deep neural networks (DNNs) and recurrent neural networks (RNNs) have significantly enhanced the accuracy of speech recognition systems, enabling better performance in transcribing speech across various languages and accents [4].

Multilingual and Cross-lingual Capabilities: DL models now support multilingual and cross-lingual speech processing, facilitating real-time translation and communication among diverse linguistic populations. This advancement promotes inclusive and equitable access to education [24].

Emotion Recognition: DL models can now analyze speech to detect emotions, providing insights into learners' engagement and motivation levels. Emotion recognition technology helps educators tailor their approaches to support students emotionally and cognitively, fostering a positive learning environment [2].

VIII. Challenges

Challenges in Developing Speech Processing Systems

Data Privacy and Security: The collection and processing of speech data raise significant privacy and security concerns. Ensuring that learners' data is protected and used ethically is crucial for maintaining trust and compliance with regulations [16].

Bias and Fairness: Speech processing systems can exhibit biases based on the data they are trained on. Addressing biases related to language, accent, gender, and ethnicity is essential to ensure fair and equitable access to educational technologies [26].

Scalability and Infrastructure: Developing scalable and efficient speech-processing systems requires robust infrastructure and computational resources. Sustainable computing practices are necessary to balance the demand for high performance with the need for energy efficiency and environmental sustainability [15].

Challenges in Achieving SDG 4

Despite significant progress, achieving SDG 4 remains a challenge due to various barriers:

Access to Education: Millions of children and adults still lack access to education, particularly in low-income countries and conflict-affected regions. [10] discussed about diabetic retinopathy

from retinal pictures utilizing cooperation and information on state of the art sign dealing with and picture preparing. The Pre-Processing stage remedies the lopsided lighting in fundus pictures and furthermore kills the fight in the picture. Although the Disease Classifier step was used to identify arising wounds and other data, the Division stage divides the image into two distinct classes. The methodology for ensuring red spots, exhausting and recognizing evidence of vein-lobby hybrid focuses was also developed in this work, using the hidden data, shape, size, object length to expansiveness distribution as contained in the general fundus picture in the problem area. Besides the Diabetic Retinopathy (DR) analysis, two graphical user interfaces (GUIs) were produced throughout this project.

Quality of Education: Many education systems struggle with providing quality education that equips learners with the skills needed for the 21st century.

Equity: There are persistent disparities in education access and outcomes based on gender, socioeconomic status, ethnicity, and disability [14].

Financing: Adequate funding is crucial for education systems to function effectively and inclusively. Many countries face financial constraints that hinder their ability to invest in education.

Unavailability of network infrastructure: There are still many places especially the rural areas where online education or accessing resources is not possible due to the unavailability of network infrastructure or facility. People in those areas are unable to take advantage of advanced learning technology tools. This reduces the quality and equity of education.

IX. Future Research Directions for Enhanced Education

Integration with Augmented Reality (AR) and Virtual Reality (VR)

Combining speech processing and DL with AR / VR technologies creates immersive and interactive learning experiences. These technologies enhance education by making content more interactive, engaging, and motivating [20,25,45], simulating real-world scenarios [46], and bridging theoretical knowledge and practical application. They also connect learners with global experts, fostering collaborations and resource exchanges [44].

Collaborative Learning Platform

Future advancements should develop collaborative learning platforms utilizing speech processing to improve communication and cooperation among learners. These platforms can facilitate group projects, peer learning, and global classrooms, fostering community and shared learning by connecting learners worldwide, thus promoting inclusivity and diversity [11].

Continual Learning and Adaptation

To stay relevant and effective, speech processing systems must be designed for continual learning and adaptation. Regular updates with new data and feedback ensure these systems meet educational needs. Integrating feedback loops, where learners' performance data refines learning algorithms, keeping educational tools aligned with current pedagogical practices [2]. Innovations like ChatGPT and Meta AI showcase how AI can continuously learn and adapt, offering personalized real-time assistance. [12] discussed that The study of viruses and their genetics has been an opportunity as well as a challenge for the scientific community. The recent ongoing SARSCov2 (Severe Acute Respiratory Syndrome) pandemic proved the unpreparedness for these situations. Not only the countermeasures for the effect caused by virus need to be tackled but the mutation taking place in the very genome of the virus is needed to be kept in check frequently.

Leveraging Unsupervised Learning Techniques

One promising direction is using unsupervised learning to minimize the need for large labeled datasets. This approach leverages abundant unlabelled speech data, enabling the creation of robust models with fewer resources. Unsupervised learning aids in developing adaptive educational tools that improve with new speech data, enhancing scalability and efficiency [38].

Enhancing Teacher Support Systems

Future research should develop advanced teacher support systems using speech processing and DL. These systems can automate administrative tasks, provide real-time feedback on teaching methods, and suggest personalized instructional strategies based on learner performance data, thereby reducing teachers' administrative burden and enhancing instructional quality.

Strategic Policy Recommendations for Maximizing Impact

Strategic policy recommendations are crucial for maximizing the impact of speech processing technologies on sustainable development. They can guide policymakers in promoting eco-friendly educational technologies and initiatives that improve digital literacy and access to education.

X. CONCLUSION

Integrating deep learning with speech processing can significantly enhance lifelong learning by making education more accessible, personalized, and engaging. Recent advancements in these technologies, along with sustainable computing practices, address current challenges and offer innovative educational solutions. This paper highlights how these technologies can support lifelong learning and quality education in the sustainable development context. They promote efficient, eco-friendly practices and continuous educational growth. As deep learning models evolve, they promise to improve education for diverse learners globally and potentially boost teaching outcomes, especially for rural educators. [47,48].

REFERENCES

[1] Y. LeCun Y, Y. Bengio & G. Hinton. Deep learning. Nature, 521(7553): 436-444, 2015.

[2] I. Goodfellow, Y. Bengio & A. Courville. Deep Learning 2016. MIT Press.

[3] A.B. Nassif, I. Shahin., I. Attili., M. Azzeh & K.Shaalan. Speech recognition using deep neural networks: A systematic review. IEEE Access, 7: 19143-19165, 2019. doi: 10.1109/ACCESS.2019.2896880.

[4] Deng L, & Li X. Machine Learning Paradigms for Speech Recognition: An Overview. IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, 21(5): 1060-1089, 2013.

[5] G. Hinton, L. Deng, D. Yu, GE. Dahl, A.R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, B. Kingsbury. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups, IEEE Signal Processing Magazine, 29(6):82-97, 2012, doi: 10.1109/MSP.2012.2205597.

[6] X. Wu, X. Zhu, G.-Q. Wu, and W. Ding. Data mining with big data. IEEE transactions on knowledge and data engineering, 26(1): 97-107,2014. doi: 10.1109/TKDE.2013.109.

[7] E. Strubell, A. Ganesh, and A. McCallum. Energy and policy considerations for deep learning in NLP. In Proc: the 57th annual meeting of the Association for Computational Linguistics, 3645-3650, 2019. doi: 10.18653/v1/P19-1355.

[8] R. Gholami, R.T. Watson, A. Molla, H. Hasan and N. Bjorn-Andersen. Information Systems Solutions for Environmental Sustainability: How Can We Do More?. Journal of the Association for Information Systems. 17(8):521-536. 2016. DOI:10.17705/1jais.00435.

[9] C.R. Aragon, S.S. Poon and A. Sears. Sustainable Software Engineering. In: Handbook of Sustainable Engineering. Pp. 665-684, 2013. Springer.

[10] Christo Ananth, D.R. Denslin Brabin, Jenifer Darling Rosita, "A Deep Learning Approach To Evaluation Of Augmented Evidence Of Diabetic Retinopathy", Turkish Journal of Physiotherapy and Rehabilitation, Volume 32,Issue 3, December 2021,pp. 11813-11817.

[11] V. Kepuska and G. Bohouta. Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In: 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, IEEE. 2018, pp. 99-103, doi: 10.1109/CCWC.2018.8301638.

[12] Christo Ananth, Pranav Pushkar, Preeti Nagrath, Jehad F. Al-Amri, Vividha, Anand Nayyar, "Mutation Prediction for Coronaviruses using Genome Sequence and Recurrent Neural Networks", Computers, Materials & Continua, Tech Science Press, Volume 73, Issue 1, April 2022,pp. 1601-1619.

[13] Y.Tonegawa. Education in SDGs: What is Inclusive and Equitable Quality Education?. In: S. Urata, K. Kuroda, Y. Tonegawa (Eds). Sustainable Development Disciplines for Humanity. Sustainable Development Goals Series. Springer, Singapore, 2023. https://doi.org/10.1007/978-981-19-4859-6_4.

[14] United Nations. The Sustainable Development Goals Report 2020. Retrieved from https://unstats.un.org/sdgs/report/2020/.

[15] S. Han, J. Pool, J. Tran and W. Dally. Learning both Weights and Connections for Efficient Neural Network. In: Advances in Neural Information Processing Systems 28 (NIPS 2015). 1135-1143,2015.

[16] W. Shi, J. Cao, Q. Zhang, Y. Li and L. Xu. Edge Computing: Vision and Challenges. IEEE Internet of Things Journal. 3(5): 637-646, 2016, doi: 10.1109/JIOT.2016.2579198.

[17] A. B. F. Mansur, N. Yusof, A. H. Basori. Personalized Learning Model based on Deep Learning Algorithm for Student Behaviour Analytic, Procedia Computer Science,163,2019, pp.125-133, https://doi.org/10.1016/j.procs.2019.12.094.

[18] A. Shemshack, J. M. A. Spector. A systematic literature review of personalized learning terms. Smart Learn. Environ. 7, 33 (2020). https://doi.org/10.1186/s40561-020-00140-9.

[19] D.L. Taylor, M. Yeung, and A.Z. Bashet. Personalized and adaptive learning. In Ryoo J.& Winkelmann K. (Eds.), Innovative learning environments in STEM higher education: Opportunities, challenges, and

looking forward. Springer Briefs in Statistics. Springer. pp. 17-34,2021. Cham. https://doi.org/10.1007/978-3-030-58948-6_2.
[20] A. Bhutoria. Personalized education and Artificial Intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model, Computers and Education: Artificial intelligence,3,2022,100068, https://doi.org/10.1016/j.caeai.2022.100068.
[21] K.J. Laak and J. Aru. AI and personalized learning: Bridging the gap with modern educational goals. 2024. https://arxiv.org/html/2404.02798v1.
[22] F.R. Sharma, and S.G. Wason. Speech recognition and synthesis tools: Assistive technology for physically disabled persons. Int. J. of computer science and Telecommun, 3(4):44-50.
[23] K. Zdravkova, V. Krasniqi, F. Dalipi and M. Ferati. Cutting-edge communication and learning assistive technologies for disabled children: An artificial intelligence perspective. Frontiers in Artificial Intelligence, 5, 970430. 2022.doi: 10.3389/frai.2022.970430. PMID: 36388402; PMCID: PMC9650429.
[24] R. Sproat, A.W. Black, S. Chen S, S. Kumar, M. Ostendorf and C. Richards. Normalization of non-standard words. Computer Speech & Language, 15(3):287-333,2001, doi:10.1006/csla.2001.0169
[25] Wang X., Young G.W., Iqbal M.Z. et al. The potential of extended reality in Rural Education's future – perspectives from rural educators. Educ Inf Technol 29, 8987–9011 (2024). https://doi.org/10.1007/s10639-023-12169-7.
[26] X. Xiong, L. Wu, F. Alleva, J. Droppo, X. Huang and A. Stolcke. The Microsoft 2017 Conversational Speech Recognition System. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 2018, pp. 5934-5938, doi: 10.1109/ICASSP.2018.8461870.
[27] T.L. Chiu, H.C. Liou, and Y. Yeh. A study of web-based oral activities enhanced by automatic speech recognition for EFL college learning. Computer Assisted Language Learning. 20(3):209-233,2007. DOI:10.1080/09588220701489374.
[28] S. O. Arik, M. Chrzanowski, A. Coates, G. Diamos, A. Gibiansky, Y. Kang, X. Li, J. Miller, A. Ng, J. Raiman, S. Sengupta and M. Shoeybi. Deep voice: Real-time neural text-to-speech. arXiv preprint arXiv:1702.07825. retrieved from https://doi.org/10.48550/arXiv.1702.07825.
[29] Y. Wang, R.J. Skerry-Ryan, D. Stanton, Y. Wu, R. J. Weiss, N. Jaittly, Z. Yang, Y. Xiao, Z. Zhu, Y. Wu, Y. Xiao and Z. Chen.Tacotron: Towards End-to-End Speech synthesis. arXiv preprint arXiv:1703.10135. Retrieved from https://arxiv.org/abs/1703.10135.
[30] T. Kabudi, I. Pappas and D.H. Olsen. AI-enabled adaptive learning systems: A systematic mapping of the literature, Computer and Education: Artificial Intelligence. 2(2021).100017. https://doi.org/10.1016/j.caeai.2021.100017.
[31] P. Muller, F. De Wet, C. VanDer Walt and T. Niesler. Automatically assessing the oral proficiency of proficient L2 speakers. In Proc: Workshop on speech and language technology for education (SLaTE), pp. 29-32, 2009.
[32] S. Crossley and D. McNamara. Applications of text analysis tools for spoken response grading. Language Learning & Technology, 17(2): 171-192,2013.
[33] Y. Wang, M. J. F. Gales, K.M. Knill, K. Kyriakopoulos, A. Malinin, R.C. Van Dalen and M. Rashid. Towards automatic assessment of spontaneous spoken English. Speech Commun. 104, pp. 47-56,2018.
[34] Z. Liu, G. Xu, T. Liu, W. Fu, Y. Qi, W. Ding, Y. Song, C. Guo, C. Kong, S. Yang, G.Y. Huang. Dolphin: a spoken language proficiency assessment system for elementary education. In proc: The web conference 2020, pp. 2641-2647, 2020. Association for Computing Machinery, New York, NY, USA, https://doi.org/10.1145/3366423.3380018.
[35] N. Iwashita, A. Brown, T. McNamara and S. O'Hagan. Assessed levels of second language speaking proficiency: How distinct?. Applied Linguistics, 20(1):24-49,2008.
[36] N. H. De Jong, M.P. Steinel, A.F. Florijn, R. Schoonen and J.H. Hulstijn. Facets of speaking proficiency. Studies in second language acquisition, 34. Pp.5-34, 2012.

[37] Y. Li, D. Chen, X. Deng. The impact of digital educational games on student's motivation for learning: The mediating effect of learning engagement and the moderating effect of the digital environment. PLoS One.19(1): e0294350, 2024. doi: 10.1371/journal.pone.0294350.
[38] S. Ravi and H. Larcochelle. Optimization as a model for few-shot learning. In Proc: International conference on learning representations (ICLR).2017.
[39] A. Haleem, M. Javaid, M.A. Qadri and R. Suman. Understanding the role of digital technologies in education: A review, Sustainable Operations and Computers, 3, pp. 275-285,2022. https://doi.org/10.1016/j.susoc.2022.05.004.
[40] J. Keengwe, M. Bhargava. Mobile learning and integration of mobile technologies in education. Educ Inf Technol 19, 737–746, 2014. https://doi.org/10.1007/s10639-013-9250-3.
[41] P. L. Rogers. Barriers to adopting emerging technologies in education. Journal of educational computing research, 22(4):455-472,2000. Retrieved from Google Scholar.
[42] S. Dreimane and R. Upenieks. Intersection of Serious Games and Learning Motivation for Medical Education: A Literature Review. In I. Management Association (Ed.), Research Anthology on Developments in Gamification and Game-Based Learning. pp. 1938-1947, 2022. IGI Global. https://doi.org/10.4018/978-1-6684-3710-0.ch093.
[43] A. Dengel. What is immersive learning?. In: 2022 8th International Conference of the Immersive Learning Research Network (iLRN). Pp. 1-5, 2022.IEEE.
[44] X. Wang, G. W. Young, A. Plechata, C. Mc Guckin and G. Makransky. Utilizing virtual reality
[45] to assist social competence education and social support for children from under-represented backgrounds. Computers & Education, 201, 104815, 2023. Elsevier. https://doi.org/10.1016/j.compedu.2023.104815.
[46] H.A.S. Garduno, M. I.E. Martinez and M. P. Castro. Impact of virtual reality on student motivation in a high school science course. Applied Sciences, 11(20), 9616, 2021, Appl. Sci. https://doi.org/10.3390/app11209516
[47] C. Andrews, M.K. Southworth, J.N. Silva and J.R. Silva. Extended reality in medical practice. Current Treatment Options in Cardiovascular Medicine, 21(4): 1-12, 2019.
[48] M.R. Blanchard, C.E. LePrevost, A. D. Tolin and K.S. Gutierrez. Investigating technology enhanced teacher Professional Development in rural, High-Poverty Middle Schools. Educational Researcher, 45(3):207-220, 2016. https://doi.org/10.3102/0013189X16644602.
[49] D. Wu, X. Yang, W. Yang, C. Lu and M. Li. Effects of teacher and school level ICT training on teachers' use of digital educational resources in rural schools in China: A multilevel moderation model. Int.J. of Educational Research, 111,101910. https://doi.org/10.1016/j.ijer.2021.101910.

AUTHOR(S) BIOGRAPHY

Ms. Rupali Khaklary is a Research Scholar pursuing a Ph.D. in the Department of Computer Science at Krishna Kanta Handiqui State Open University, Guwahati, Assam. Her research interests are in the fields of speech processing and machine learning. Alongside her studies, she is currently serving as a Guest Faculty at Kokrajhar Govt. College, Kokrajhar, Assam, where she contributes to the academic development of students in the field of computer science.