UKLO Annual International Scientific Conference 2024 • Proceedings
Published by the University "St. Kliment Ohridski" – Bitola

(*) Corresponding author
Original research article
DOI: 10.20544/AISC.1.1.25.P18

#UKLO Annual International Scientific Conference 2024

USING AUTOMATIC SPEECH RECOGNITION TO SUPPORT STUDENTS WITH DISABILITIES

Aleksandar Kotevski*
email: kotevski.ace@gmail.com

Abstract

The transformative potential of digital technologies, such as using the benefits of AI and automatic speech recognition techniques, in order to improve the accessibility of students with disabilities in the education, is the main focus of this paper. As education increasingly shifts toward digital platforms, traditional barriers can be mitigated through innovative tools that cater to diverse needs. Automatic speech recognition (ASR) techniques empowers students with disabilities by providing real-time transcription, facilitating note-taking, and enhancing participation in classroom discussions. Additionally, it offers immediate feedback for language learners and supports those with writing difficulties, allowing for the expression of ideas in a verbal format. The goal of this paper is to propose an approach for implementing ASR techniques in combination with NLP translation in the educational process for students with disabilities. The proposed approach aims to create inclusive and effective learning environment for students with disabilities. Comparative analysis of several AI pre-trained models available on Hugging Face, including Wav2Vec 2.0, HuBERT, DeepSpeech and Jasper is covered in the paper. Furthermore, main challenges and best practices for implementing the proposed approach are included too.

Keywords: AI, ASR, NLP, disabilities, education, learning, improvements

INTRODUCTION

The main goal of inclusive education is to ensure that students with disability will have access to the required learning materials. Many students continue to face barriers in accessing educational content, particularly in traditional classroom settings. By enhancing communication, personalizing learning experiences and automating accessibility features, AI and ASR technologies have the potential to significantly improve educational outcomes for students with disabilities.

Students with disabilities represent a significant portion of the student population worldwide. According to the World Health Organization (WHO), approximately 15% of the global population lives with some form of disability, many of whom are students.

The need for more adaptive and technology-driven solutions to overcome these challenges and limitations led to the increasing use of AI and ASR technologies in educational contexts.

In the era of technology revolution and implementing AI everywhere, we need to take it in consideration and to focus to the transformative potential for educational technology. AI tools can assist in various areas, from adaptive learning platforms to content generation and personalized assessment tools. For students with disabilities, AI can provide customized

support that adapts to individual needs in format adequate to the student needs and requirements.

Over the past several years, AI technologies introduced a lot of changes in the field of education. In that content, AI has been particularly beneficial in supporting students with disabilities by offering solutions to make educational environments more inclusive. Speech recognition systems (ASR) and text-to-speech tools help students with hearing or visual impairments. Furthermore, AI-powered tutoring systems can assist those with learning disabilities by providing additional support and guidance tailored to their specific challenges.

RELATED PAPERS

The impact of AI in special need education to promote inclusive pedagogy (Shalini, 2020) has been analyzed from the authors, based on the data conducted using focused interviews from teachers and students with disabilities. Also, they proposed a framework for an inclusive future of Special Need Education based on focused interview. According to authors (Drigas and Ioannidou, 2013), AI has the ability to score essays automatically, determine the reading and writing challenges of SWLDPs, develop psychological profiles for SWLDPs, and assess their spelling issues.AI enabled computer vision systems can detect a variety of physical behaviors, including emotional expressions and gestures, which can be beneficial for students with autism or emotional impairments (Asthana & Gupta, 2019). A study by Kukulska-Hulme et al. (2021) found that AI-driven platforms significantly improved engagement and academic outcomes for students with learning disabilities by offering adaptive reading materials and alternative learning formats, such as interactive simulations and audio-based content. A study by Johnson & Patel (2022) demonstrated that AI-based captioning systems increased the accessibility of online lectures for students with hearing impairments, allowing them to follow content more effectively.

In other paper, the authors (Waterfield, D. A., Watson, L., & Day, J. 2024) compared five existing platforms which have features dedicated to the students with disabilities. They describe how each AI platform can be utilized by special education teachers within their practice is conveyed. Overall implications regarding AI integration, usage, and ethical considerations in special education practice are discussed.

In the next related paper, the author (Shipra, Sharma 2020) described some advancements of artificial intelligence and its use in our daily life, followed by discussing the challenges faced by the teacher in classroom pedagogy, challenges faced by students with disabilities in communication, and access to various educational resources in the process of inclusion in mainstream institutions. Sdenka Zobeida completed systematic review analyzes the advantages and challenges of using AI and new technologies in different sociocultural contexts, and their impact on minority students. In terms of advantages, they found that AI and new technologies improved student performance, encouraged student interest, promoted student engagement and showed other advantages. This review also identifies the main challenges associated with the use of AI. According to Chung et al. (2021), AI-based ASR systems are increasingly being used to provide captions for lectures and group discussions. These systems offer high accuracy, even in noisy classroom environments, and allow students to access spoken content that they might otherwise miss. Patel et al. (2020) found that these systems significantly improved communication for students with cerebral palsy, enabling them to interact more freely in the classroom.

Li & Wang (2022) highlighted the effectiveness of AI-powered language learning tools, which can provide personalized practice based on the student's individual needs. The authors (Abbasi, Zulfiqar, Rasool and Quadri, 2024) in another related paper are focused on several use-case scenarios of implementing AI in the educational process for the learners with

disability, pointing the benefits of using those technologies. The impact of AI technologies to the educational process for the learning with different type of disability covered by Dr. Smita Tiwary Ojha, 2022.

A case study conducted at the University of California, Berkeley, integrated an AI-powered ASR system to support students with hearing impairments. The system transcribed lectures in real time with over 95% accuracy, allowing students to access spoken content on their devices immediately. Students reported a 30% improvement in lecture comprehension and participation in class discussions (Smith et al., 2023).

In a study at a public middle school in London, an adaptive learning platform using AI was introduced to support students with dyslexia. The platform used AI to analyze each student's reading ability and adjusted the difficulty of the text accordingly. Students showed an average improvement of 40% in reading fluency and comprehension over a 12-week period (Gordon & Green, 2021).

A pilot program at a specialized school for students with speech impairments utilized an ASR system to aid speech therapy. The system was trained on the speech patterns of students with speech disorders and provided real-time feedback. 80% of students in the program showed improvements in speech clarity and confidence in communication within six months of using the system (Lee & McKinnon, 2022).

HUGGING FACE MODEL FOR ASR

Hugging Face providing tools, models, and infrastructure to help streamline the development, deployment, and collaboration of different machine learning solutions. With its open-source orientation and focus on usability, Hugging Face has become a go-to platform for both beginners and advanced practitioners in the AI field. Their tools, such as the Transformers library, provide high-level APIs that allow users to quickly implement and deploy AI models.

In that content, there are several pre-trained models for ASR. The following table shows a comparative analysis of the most used pre-trained models short overview of them, their characteristics and use case scenarios.

Table 1: comparative analysis

Model	Features	Use Case
Wav2Vec 2.0	- Self-supervised learning - High accuracy in noisy environments - Fine-tuning on labeled data - Works in multiple languages	General ASR tasks, transcription, low-resource language support, noisy environments
HuBERT	- Self-supervised learning - Robust to noise - Fine-tuned for domain-specific tasks - Multilingual support	Domain-specific ASR, noisy environments, fine-tuning for specialized use cases
DeepSpeech	- Lightweight and efficient - Real-time transcription - Open-source - Lower latency	Real-time ASR, open-source projects, low-latency speech recognition
Whisper	- Multilingual support (100+ languages) - Transcription and translation - Robust to accents and background noise - Fine-tuning for specific	Multilingual ASR, translation from speech, transcription of noisy or accented speech

	domains	
Kaldi	- Highly customizable - Supports multiple ASR pipelines - Domain-specific fine-tuning (e.g., medical or legal) - Strong performance in industrial settings	High-accuracy ASR for specialized domains, research and custom ASR pipelines
Conformer	- Combines convolution layers with transformers - Optimized for ASR - Effective in noisy conditions - High accuracy and low latency	Speech recognition in noisy environments, real-time applications
VQ-Wav2Vec	- Combines vector quantization with Wav2Vec - Faster fine-tuning - High scalability - Efficient for large-scale deployment	Scalable and efficient ASR, low-latency transcription, low-resource languages

NATURAL LANGUAGE PROCESSING FOR TRANSLATION

The use of Natural Language Processing (NLP) for translation has opened up exciting possibilities for increasing access to educational materials, particularly for students with disabilities or those who speak different languages.

The Helsinki NLP team has significantly advanced the field of neural machine translation (NMT) by developing a series of pre-trained translation models available on Hugging Face. These models, based on the OPUS dataset, support a wide range of language pairs and demonstrate high precision in translation tasks. Currently, they support above 90 natural languages.

PROPOSED APPROACH

If we combine the benefits and features of ASR and NLP translation technologies, we can extend the range of accessibility of audio and video materials for the students. That approach can bring a lot of benefits for the students with disabilities.

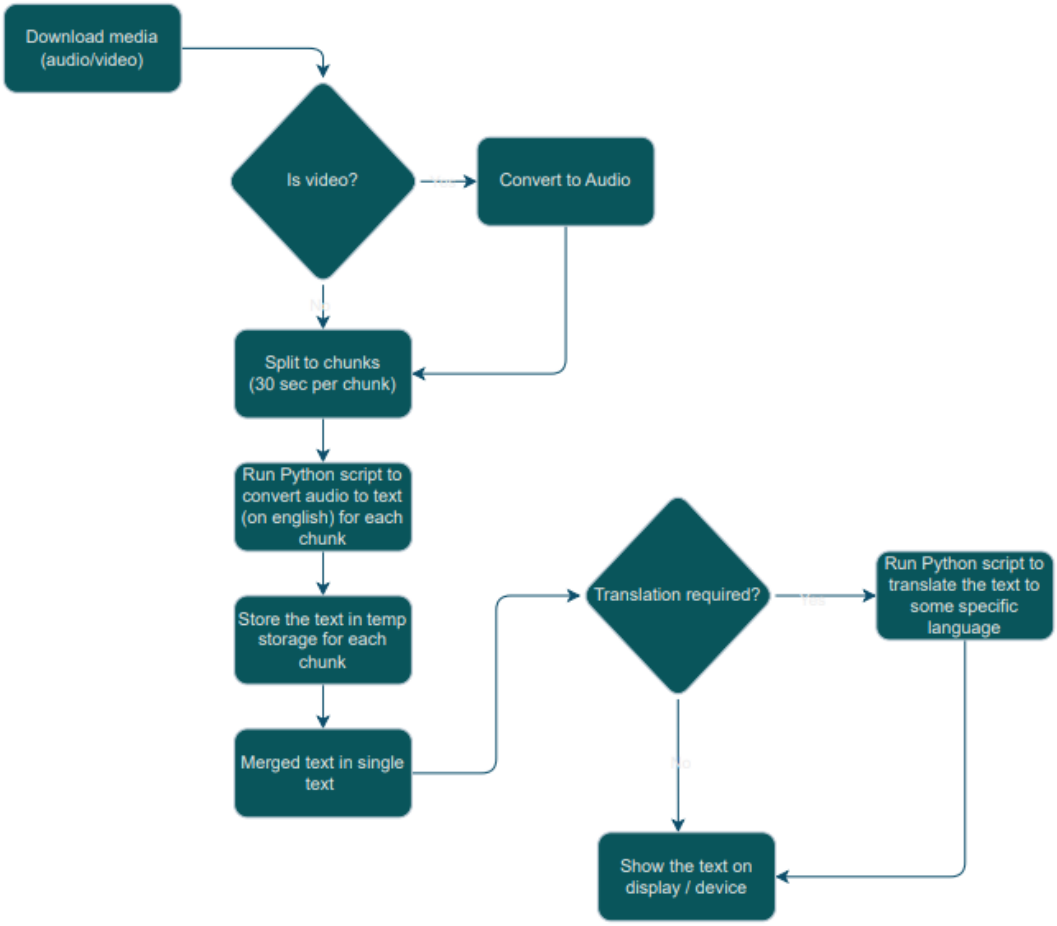

Figure 1: Flow diagram

In the scope of this paper, we suggest a combination of ASR and NLP translation. Based on that, student with disability can get transcription of audio and video materials. Furthermore, the students can get the transcription in the other natural language (any natural language which is supported by NLP Helsinki.

As an input, the system expects video or audio materials. Because ASR expects audio, we need to convert the video material to audio (for video inputs). In our research, we used python package AudioSegment from pydub to extract the audio from the video materials. After converting to audio, we are splitting input in smaller chunks and run the service for transcription to English for each chunk. Chunks can we process parallel, on the same machine or different machines, based on the available hardware. When all chunks will be handled, we are merging the results in single text.

Then, we can translate that text in other natural language by using Helsinki NLP.

CONCLUSION

ASR has been recognized as a key technology for enhancing the educational experience for students with hearing disabilities. They convert spoken language into written text in real-time, which is invaluable for the students. The integration of AI and ASR technologies into educational environments offers significant potential for supporting students with disabilities. These technologies can personalize learning, enhance communication, and increase accessibility for students with disability. However, their

successful implementation depends from several aspects, such as adequate hardware and teachers training.

With continued innovation and the proper support structures, AI and ASR can play a pivotal role in creating more inclusive and accessible educational experiences for all students. Although AI and ASR technologies hold great potential, their implementation in educational settings is not without challenges, such as technical limitations and the need for teacher training.

The use of AI and ASR involves the collection and processing of sensitive data, such as speech patterns, learning progress, and behavioral data. That's why, data security and privacy issues must be addressed to ensure that student information is protected in compliance with regulations.

On the other hand, continuous efforts are needed to train AI systems on diverse datasets to minimize the issues related with inaccuracies and bad transcription, mainly due to trained models with audio from people with non-standard accents or dialect.

The high cost of required hardware is a significant barrier when implementing ASRin the educational process. While ASR can greatly enhance accessibility for students by providing tools like real-time transcription, voice commands, and language support, the need for powerful GPUs / CPUs to train and run sophisticated ASR models can make these technologies financially inaccessible to many educational institutions.

Future more, there are additional expenses, including the costs for power consumption, cooling systems, hardware repairs and etc.

The proposed approach has several challenges. Audio material length - required time for transcription depends from the input length. Means that more processing time will be required for the longer input materials. This challenge can be resolved by splitting the input in more parts, and handle each part of separate machine. Then all machines will finish with the processing, the results can be merged in single transcription. Different sound quality – bad audio quality, noise or some specific language dialects can affect the quality of the transcription. At this moment, for some natural language, they are learning models with small amount of training data, which leads to not very precise transcription. Also, the expensive hardware and equipment can be a real challenge for some educational institution and limit them to implement the proposed approach in the educational process.

To achieve better result, as per our testing, we are suggesting using not very large audio materials. Also, we suggest transcription to English first, and then to use Helsinki NLP to translate the transcript to other natural language versus directly transcription the input material to some non-English natural language, because the trained models for English contains much more trained datasets.

REFERENCES

1. Shalini Garg, Impact of Artificial Intelligence in Special Need Education to Promote [2] Inclusive Pedagogy, International Journal of Information and Education Technology, July 2020
2. Drigas, A., &Ioannidou, R.-E. (2013). A review on artificial intelligence in special education. Communications in Computer and Information Science, 385–391.
3. Asthana, P., & Gupta, V. K. (2019). Role of artificial intelligence in dealing with emotional and behavioral disorders. International Journal of Management, IT & Engineering, 9(5), 398–403.
4. Waterfield, D. A., Watson, L., & Day, J. (2024). Applying Artificial Intelligence in Special Education: Exploring Availability and Functionality of AI Platforms for Special Educators. Journal of Special Education Technology, 39(3), 448-454. https://doi.org/10.1177/01626434241257237

5. Shipra Sharma, Impact of Artificial Intelligence in Special Need Education to Promote Inclusive Pedagogy, International Journal of Information and Education Technology, 2020
6. Sdenka Zobeida, Artificial Intelligence and New Technologies in Inclusive Education for Minority Students: A Systematic Review
7. Dr. Muhammad Umar Riaz Abbasi, Anum Zulfiqar, Muhammad Shoaib Rasool, & Dr. Syed Shameel Ahmed Quadri. (2024). Impact of AI on the inclusion of Learners with Special needs: Public Policy Perspective in Contemporary Scenario. Arbor. https://doi.org/10.1234.arbor.1`1.1.05
8. Dr. Smita Tiwary Ojha, Artificial Intelligence In Special Education, Id&Cp, Journal of Positive School Psychology, 2022
9. Kukulska-Hulme, A., et al. (2021). "AI in Education: Personalizing Learning for Students with Disabilities." *Journal of Educational Technology & Society*, 24(2), 85-97. Johnson, M., & Patel, R. (2022). "Automated Captioning and Its Impact on Accessibility in Online Education." *Educational Technology Research and Development*, 70(1), 21-36. Chung, S., et al. (2021). "Real-Time Speech-to-Text Applications in Educational Settings: Enhancing Access for Deaf Students." *Journal of Educational Accessibility*, 15(3), 128-144.
10. Li, Y., & Wang, H. (2022). "Leveraging AI for Language Learning: Applications for Dyslexic Students." *International Journal of Educational Technology*, 30(2), 92-106.
11. Smith, J., et al. (2023). "The Effectiveness of AI-Powered ASR Systems in Higher Education for Deaf Students." *Journal of Higher Education Accessibility*, 39(1), 101-118.
12. Gordon, L., & Green, D. (2021). "Adaptive Learning Platforms for Students with Dyslexia: A Case Study." *Journal of Special Education Technology*, 34(4), 78-94
13. Lee, T., & McKinnon, A. (2022). "Speech Therapy with ASR: A Pilot Study on Improving Communication for Students with Speech Impairments." *Journal of Speech and Language Pathology*, 19(3), 67-82.
14. https://blogs.helsinki.fi/language-technology