International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

SOFTWARE FOR SPEECH LEARNING THERAPY

[1]Dr. M. Ratna Raju,[2]B. Thirupathamma,[3]Ch. Gowtham Manikanta,[4]A. Sammer Babu

[1]Associate Professor, Dept of computer Science and Engineering, St. Ann's College of Engineering and Technology, Chirala-523187, India.

[2,3,4]U. G Student, Dept of Computer Science and Engineering, St. Ann's College of Engineering and Technology, Chirala-523187, India.

ABSTRACT-This project presents an Artificial Intelligence (AI)-based Software for Speech Learning Therapy designed to improve speech pronunciation, fluency, articulation, and communication skills through intelligent speech analysis. The proposed system integrates Automatic Speech Recognition, PyAudio, HTML, CSS, JavaScript, and JSON to provide real-time speech assessment and personalized learning support. Users practice randomly generated sentences through a microphone, and the recognized speech is compared with the original text to evaluate pronunciation accuracy, fluency, speech proficiency, and incorrectly spoken words. The application delivers instant feedback and maintains user performance records for continuous progress monitor the practical application of Artificial Intelligence and self-learning. The intelligent speech evaluation speech therapy by enabling users to practice independently at any time. The proposed system demonstrates the practical application of Artificial Intelligence in speech learning and educational technology, improving user engagement, communication skills, and overall learning efficiency while providing a scalable platform for future intelligent speech therapy applications.

KEYWORDS: Artificial Intelligence (AI), Speech Learning Therapy, Automatic Speech Recognition (ASR), Speech-to-Text (STT), Pronunciation Assessment, Fluency Analysis, Python, Flask, Speech Recognition, Real-Time feedback.

INTRODUCTION

Speech Learning Therapy aims to improve pronunciation, fluency, articulation, and communication skills through intelligent speech analysis. With rapid advancements in Artificial Intelligence (AI) and

Automatic speech Recognition (ASR), speech evaluation has become more accurate, efficient, and accessible. This project combines AI-based speech recognition with an intelligent web application to provide real-time pronunciation assessment and personalized learning but also promotes continuous self-improvement by delivering instant feedback and interactive speech practice.

LITERATURE SURVEY

Speech learning and therapy systems have gained significant attention in recent years due to rapid advancements in Artificial Intelligence (AI) and Automatic Speech Recognition (ASR) technologies. Speech Learning Therapy is designed to improve pronunciation, fluency, articulation, and communication skills through intelligent speech analysis and real-time feedback. The core concept of speech learning systems revolves around speech recognition, accessible, and effective platform for continuous speech improvement.

LITERATURE SURVEY

Speech learning and therapy systems have gained significant attention in recent years due to rapid advancements in Artificial Intelligence (AI) and automatic Speech Recognition (ASR) Technologies. Speech Learning Therapy is generally designed to improve pronunciation, fluency, articulation, and communication skills through intelligent speech learning systems resolves around speech recognition, AI-based evaluation, and personalized learning support to provide an interactive, accessible, and effective platform for continuous speech improvement.

RELATED WORK

Several researchers have contributed to the development of AI-based speech learning and therapy systems with the aim of improving pronunciation, fluency, articulation, and communication skills. Initial research in this domain focused on traditional speech therapy methods, where learners received guidance through face-to-face sessions with speech therapists. These approaches relied on manual assessment and regular supervision, offering limited accessibility, higher costs, and reduced flexibility for continuous speech practice.

EXISTING SYSTEM

The existing speech learning and therapy systems mostly depend on traditional teaching methods, manual pronunciation assessment, or therapist-guided sessions. Users must attend regular training sessions to receive feedback on their speech and communication skills. Speech evaluation is

often performed manually, making the process time-consuming and dependent on the availability of speech therapists. Most existing systems provide limited real-time feedback and lack intelligent mechanisms for automatic pronunciation assessment and continuous assessment and continuous progress monitoring. Overall, the existing system lacks accessibility, intelligent speech analysis, personalized learning, and cost-effective continuous speech practice.

PROPOSED SYSTEM

The proposed system implements an **AI-based Speech Learning Therapy** application for intelligent speech analysis and pronunciation assessment. It allows users to practice randomly generated sentences using a microphone and evaluates their speech in real time. The system compares the organized Speech with the original sentence to calculate pronunciation accuracy, fluency, speech level, and incorrectly spoken words. This approach significantly improves speech learning by providing instant feedback and personalized practice. The system eliminates the need for continuous therapist supervision, making speech learning more accessible and cost-effective. User information and learning progress are securely stored in JSON database for continuous monitoring and evaluation.

SYSTEM ARCHITECTURE

The architecture of the AI-based Speech Learning Therapy system consists of a **User Interface, Speech Recognition Module, AI Evaluation Engine, JSON Database, and Feedback Module**. The user speaks the displayed sentence through a microphone, and the Speech Recognition module converts the spoken input into text.

The AI Evaluation Engine compares the recognized text with the original sentence and calculates pronunciation accuracy, fluency, speech level, and incorrectly spoken words. Based on the evaluation results, the Feedback Module displays instant feedback to the user, while the JSON database securely stores user information and learning progress. This architecture ensures accurate, interactive, and real-time speech learning and pronunciation assessment.

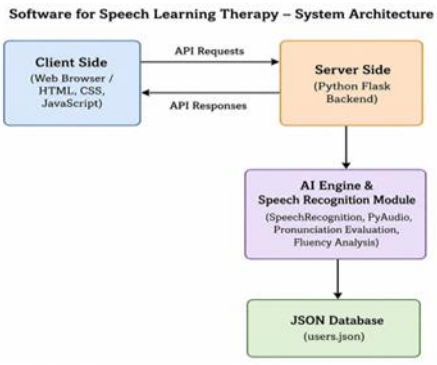

Fig.1: System Architecture

METHODOLOGY DESCRIPTION

The methodology of the Speech Learning Therapy system is designed to provide intelligent speech learning, pronunciation assessment, and continuous progress monitoring using Artificial Intelligence and Speech Recognition technologies. The system follows a structured approach that includes speech input, speech recognition, AI-based evaluation, feedback generation, and progress storage.

`RESULTS AND DISCUSSION

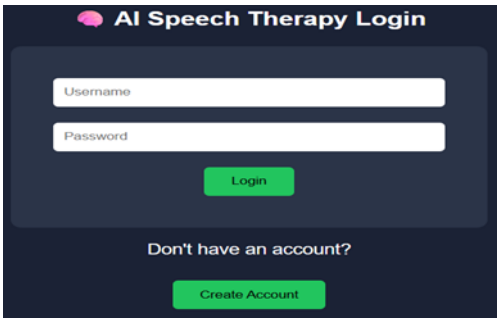

Fig.2: Login Page

The Login page serves as the main entry point of the Speech Learning Therapy application. It provides secure user authentication and allows registered users to access speech learning and pronunciation evaluation features.

Fig.3: Registration Page

The Registration page allows new users to create an account by entering their details. The user information is securely stored in the JSON database for future authentication and personalized speech learning sessions.

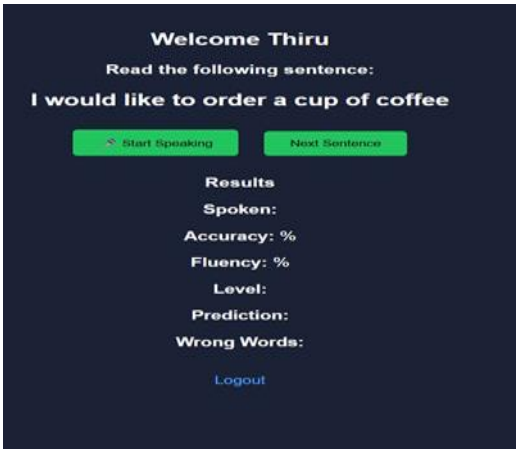

Fig.4: Home Page

The Home page displays randomly generated practice sentences for users to read aloud using the microphone. It provides easy navigation to speech practice and displays the learning interface.

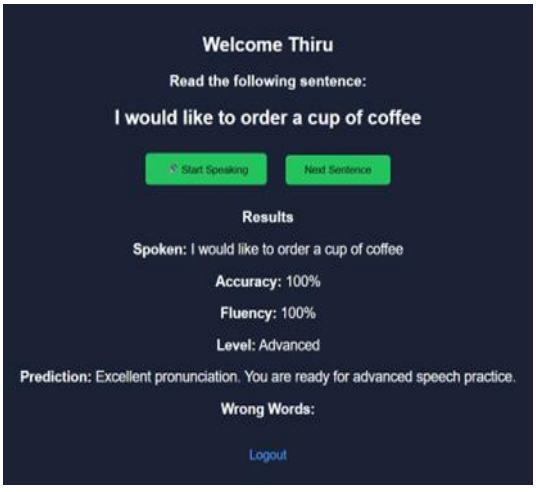

Fig.5: Speech Evaluation Page

The Speech Evaluation page is the core module of the application. It captures the

user's speech, coverts it into text, compares it with the original sentence, and displays pronunciation accuracy,

fluency, speech level, prediction, and incorrectly spoken words as real-time feedback.

CONCLUSION

The AI-based Speech Learning Therapy system enhances speech learning by providing intelligent pronunciation assessment and real-time feedback. By analyzing spoken input and evaluating pronunciation accuracy and fluency, the system offers an interactive, reliable, and cost-effective platform for improving communication skills through continuous practice.

FUTURE SCOPE

The proposed AI-based Speech Learning Therapy system can be further enhanced by integrating advanced Artificial Intelligence and Deep Learning models to improve speech recognition accuracy under different speaking conditions. The system can be extended to support multiple languages, regional, accents, emotion detection, and personalized learning recommendations. Integration with cloud storage and mobile applications can enable remote access and progress synchronization. Additional features such as therapist, dashboards, voice history, adaptive learning exercises, and offline speech analysis can further improve the effectiveness and accessibility of the system.

REFERNCES

[1] D. P. Harini, "Electronic Health Record Stage Identification using Principal Component Analysis," *Machine Intelligence Research,* vol. 18, no. 01, pp. 864- 873, 17 8 2024.

[2] D. P. Harini, "Analyze and Predict of Human Cyber Attackers using Artificial Neural Network," *Journal of Basic Science and Engineering,* vol. 21, no. 1, pp. 1529 - 1536, 8 7 2024.

[3] D. P. Harini, "Scalable and Secure Sharing of Client Medical Records in Cloud," *International Journal Of Computer Sc,* vol. 4, no. 4, pp. 161-163, 2013.

[4] Miguel Grinberg. Flask Web Development. O'Reilly Media.

[5] Eric Matthes. Python Crash Course. No Starch Press.

[6] Lawrence Rabiner and Bitting-Hwang Juang, Fundamentals of speech Recognition, Prentice Hall, 1993.

[7] Xuedong Huang, Alex Acero, and Hsiao-Wuan Hon, Spoken Language Processing: A Guide to Theory, Algorithm,

and system Development, Prentice Hall, 2001.

[8] Li Deng and Douglas O'Shaughnessy, Speech Processing: A Dynamic and Optimization-O

[9] Florian Eyben, Real-Time Speech and Music Classification by Large Audio Feature Space Extraction, Springer, 2016.

[10] Christopher D. Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press, 1999.

[11] D. O'Shaughnessy, Speech Communications: Human and Machine, @nd Edition, IEEE Press, 2000.

[12] Ben Gold and Nelson Morgan, Speech and Audio Signal Processing: Processing and Music, John Willey & Sons, 2000.

[13] Lawrence R. Rabiner and Ronald W. Schafer, Digital Processing of Speech Signals, Prentice Hall, 1978.

[14] Thomos F. Quatieri, Discrete-Time Speech Signal Processing: Principles and Practice, Pearson Education, 2002.

[15] Douglas O'Shaughnessy, speech Communication: Human and Machine, Universities Press, 2001.