Arigala Adarsh

8801466582 arigalaadarsh780@gmail.com linkedin.com/arigala-adarsh /ArigalaAdarsh

Career Objective

Research-oriented Computer Science graduate (2025) with approximately 2 years of experience in Speech and Audio AI, including a long-term research internship (May 2024 – Aug 2025) followed by a full-time research role. Seeking to contribute to impactful research and development in speech synthesis, sound understanding, and multilingual audio modeling. Passionate about building robust, low-resource and cross-lingual systems that bridge the gap between research and real-world applications.

Research Areas

Speech and Audio: Speech Synthesis (Text-to-Speech), Sound Event Detection (SED), Targeted Sound Detection (TSD), Audio Captioning (AAC), Low-Resource and Multilingual Modeling
Computer Vision: Medical Image Analysis

Education

RGUKT RK VALLEY Graduation Year 2025
Bachelor of Technology in Computer Science and Engineering (CGPA: 9.1 / 10.00) *Andhra Pradesh*

RGUKT RK VALLEY 12^{th} (2019-2021)
MBIPC (CGPA: 9.85 / 10.00) *Andhra Pradesh*

RPBS ZPHS 10^{th}(2018-2019
(CGPA: 10.00 / 10.00) *Andhra Pradesh*

Experience

Project Associate, IIT Madras October 2025 – Present
SPRING Lab, Department of Computer Science and Engineering *Chennai*

- Working under Prof. S. Umesh on multilingual and low-resource Text-to-Speech (TTS) and Voice Conversion systems.
- Developing and optimizing diffusion-based TTS models for Indian languages under the *Speech Technologies in Indian Languages* project.
- Analyzing and modifying the F5-TTS architecture for low-resource multilingual adaptation, with a focus on improving duration modeling and alignment stability.

Research Intern, IIT Hyderabad May 2024 – Sep 2025
Department of Electrical Engineering *Hyderabad*

- Worked under Prof. Sri Rama Murty in the Speech Information Processing Lab on speech and audio intelligence systems using Robot Operating System (ROS).
- Developed models for audio tagging, sound event classification, and audio search in dynamic environments.
- Proposed a unified encoder architecture for reference-guided targeted sound detection in audio mixtures, enabling reliable detection of specific target sounds in complex acoustic environments.
- Designed a custom lightweight decoder architecture for audio captioning and incorporated AudioSet-based semantic representations to improve caption generation performance.

Data Science Intern, IIM Kozhikode May 2024 – June 2024
Under Prof. Salman's supervision *Remote*

- Conducting web scraping from diverse online sources and performing meticulous data pre-processing to ensure the accuracy and reliability of collected data. Employing advanced topic modeling techniques to uncover underlying patterns and insights, contributing significantly to academic research advancement.
- Utilizing regression analysis to explore relationships and predict outcomes, providing valuable insights for decision-making processes.

Technical Skills

Programming:	Python
Deep Learning:	PyTorch, TensorFlow
Speech & Audio:	ESPnet, Torchaudio, Whisper, Librosa
ML & Data:	Scikit-learn, NumPy, Pandas
Systems:	Linux, Git, Multi-GPU Training, ROS2 (Robot Operating System 2)
Web Technologies:	HTML, CSS, Bootstrap, JavaScript, WordPress
Databases:	MySQL, Microsoft SQL Server

Projects

Automatic Audio Captioning using PANNs CNN14, ConvNeXt, and BART

- Built an Automatic Audio Captioning (AAC) system using PANNs CNN14 and ConvNeXt as audio encoders and BART as the transformer-based decoder.
- Processed raw audio into log-mel spectrograms and trained the system on Clotho, AudioCaps, and WaveCaps datasets for diverse caption learning.
- Enhanced BLEU score through effective model fine-tuning on multi-dataset training, improving the quality and relevance of generated audio captions.

Sound Event Detection using PANN CNN14

- Fine-tuned the PANN CNN14 model for sound event detection across 18 classes, achieving high accuracy in identifying various audio events.
- Integrated the model with Robot Operating System (ROS) for real-time performance, enabling seamless audio classification and interaction in dynamic environments.
- Utilized advanced techniques, such as data augmentation and feature extraction, to enhance the model's robustness and performance in diverse acoustic conditions.

Mental Health Assisting Chatbot - Providing Depression Assistance to Students

- Developed an interactive chatbot using Streamlit, designed to assist students dealing with depression by providing resources and support through conversation.
- Implemented NLP techniques for understanding user queries and providing personalized suggestions, while incorporating a depression assessment tool to evaluate user sentiment and recommend appropriate resources.

Smart Attendance Management System

- Designed and developed a web-based attendance system with a user-friendly interface for seamless record management and reporting.
- Built using HTML, CSS, Bootstrap, JavaScript, Flask, and MS SQL Server to ensure responsive design and robust backend integration.

Stress Analysis

- Collected data from approximately 800 participants in my surroundings for stress analysis, focusing on identifying factors contributing to stress and understanding stress patterns across different age categories.
- Conducted comprehensive data analysis to identify key stressors and their impact, determining which age groups were most affected and exploring changes in habits and behaviors associated with stress.

Publications

Reference-Guided Targeted Sound Detection Using a Unified Encoder
Shubham Gupta[*], **Adarsh Arigala**[*], B. R. Dilleswari, Sri Rama Murty Kodukula
In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2026*.

Parasitic Egg Detection and Classification: An Overview of Recent Progress and New Challenges
D. Gnanavenkata Kumar[*], **Adarsh Arigala**[*], Reddy Palli Trisha, Penugonda Ravikumar
In *Proceedings of SOMET 2025*.

Enhancing Audio Captioning with Auxiliary AudioSet Semantics
Shubham Gupta[*], **Adarsh Arigala**[*], Sri Rama Murty Kodukula
Manuscript under review.

[*]Equal contribution.