Openings at Prosical Language Learning Solutions

Prosical Language Learning, an IIT Bombay incubated start-up, has developed powerful tools for the assessment of oral reading fluency, reading comprehension and foundational literacy skills. The language pedagogy-informed Al-based solution uses voice to obtain reading progress data that can drive instruction in real-time. Our flagship product TARA has been used for online benchmarking and progress monitoring of oral reading fluency for over one million students across Grades 3-8. We are now poised to expand our solutions to power reading and speaking instruction while we target several Indian languages and English. We are on the lookout for Al/ML Researchers and an Audio R&D Head to lead the research and development for state-of-the-art technologies to deploy at scale for our language learning products co-designed with educators.

If you fit the roles described below, and the idea of building far-reaching impactful solutions excites you, please contact: Prof. Preeti Rao, Dept. of Electrical Engg., I.I.T. Bombay (prao@iitb.ac.in)

Audio R&D Head, AI/ML Researcher

Key Responsibilities

- 1. Develop, train, benchmark, profile and deploy Automatic Speech Recognition (ASR), Text-to-Speech (TTS), Natural Language Understanding (NLU) systems in language learning products designed to improve Reading, Speaking, and Listening skills.
- 2. Manage and maintain the annotation and processing of large speech datasets.
- 3. Collaborate with language and pedagogy experts to advance the performance of Al systems to match human experts in standardised and formative assessments.
- 4. Collaborate with the product team for system deployment in real-world applications.

Minimum Qualifications/Skills

- Ph.D./Masters in Electrical Engineering, Computer Science or AI/ML from a reputed institute with a specialization in Audio/Speech Processing
- Published research in leading conferences (Interspeech, ICASSP, etc.) / journals.
- Experience in at least one of the following
 - Hybrid Speech Recognition Systems such as Kaldi TDNN and/or End-to-End Speech Recognition Systems like wav2vec, Whisper, Conformer, DeepSpeech.
 - Experience with training and fine-tuning TDNN models in Kaldi
 - Experience with fine-tuning Wav2vec2/Whisper models
 - Experience with developing data augmentation/transfer learning techniques to improve system performance for low resource Indic language scenarios.

- Experience with audio-preprocessing to improve system performance in noisy recording conditions
- o TTS systems like (Wavenet, Tacotron, FastSpeech, VITS, StyleTTS)
 - Specific tasks related to improving naturalness and intelligibility (model training, multi-lingual, prosody/emotion modelling) associated with TTS systems for different languages
- Natural Language Processing
 - Knowledge of linguistics, and hands-on experience with LLM and Gen Al systems for natural language understanding tasks including translation.
- Strong programming experience in Python
- Experience with Machine Learning toolkits like Scikit Learn and Deep Learning frameworks such as TensorFlow or PyTorch.
- Working knowledge of Git version control systems

Preferred Qualifications/Skills

- 1-2 Years relevant work experience
- Experience with inference optimization frameworks for mobile applications like PyTorch Mobile (ExecuTorch)/ONNX
- Knowledge of C/C++