Piyush Singh Pasi

☐ (+91) 8652226601 | ✉ piyush.singh.pasi@gmail.com | ⌂ Piyush-Pasi | ⓘ piyush-pp

I enjoy working at the intersection of multimodal and multilingual models with my interests skewed toward speech and NLP.

Examination	University	Institute	Year	CPI (/10)
Post Graduation	IIT Bombay	IIT Bombay	2023	9.67
Graduation	University of Mumbai	MCT Rajiv Gandhi Institute of Technology	2020	9.27

M.TECH PROJECT & SEMINAR

- **Aligning Questions with Long Audio Survey Interviews**
  Accepted in IJCAI'23, AI for Social Good Track *(Jul'22-Mar'23)*
  (Guides: Prof. Preethi Jyothi, Prof. Ganesh Ramakrishnan)
  - **Objective**: Given a long-audio interview and a question from a fixed questionnaire, find weak temporal boundaries for the question in the interview.
  - Challenges include noisy survey data, Bihari dialect influence, on avg. **40-min. long** interviews, the spoken question in the audio is an interpretation of the question in the questionnaire.
  - Trained **End-to-End** PyTorch model in a **weakly-supervised setting.**
  - Achieved **61.8%** and **79.4%** Recall@10 for model trained with speech and it's corresponding ASR respectively. Inference supports many **unseen languages** for questions. Saves manual search in 40-min. long audio for the survey organization.

- **Multimodal Learning for Video Understanding** *(Jan'22-May'22)*
  (M.Tech Seminar, Guides: Prof. Preethi Jyothi, Prof. Ganesh Ramakrishnan)
  - Carried out a survey of papers on **Weakly-Supervised Audio-Visual Video Parsing.**
  - The survey encompasses techniques like **Modality-Specific Label Completion, Cross-Video & Cross-Modality Signals,** and **Multimodal Representations using MLM.**

RESEARCH PROJECTS

- **Audio Visual Video Parsing** *(Jan'22-May'22)*
  (Guides: Prof. Preethi Jyothi, Prof. Ganesh Ramakrishnan)
  - **Objective:** The task is to parse a video to detect audio-visual events with temporal boundaries.
  - The challenge of the task is training in **Weakly-Supervised setting** with only event labels available as a bag of labels for each video and no temporal or event modality annotations.
  - Achieved abs. gain of **2%** & **1.6%** in F1-scores of visual & audio-visual events with our improved Hybrid Attention Network.
  - Identified **weak spot** caused by **label smoothing** technique on Attentive MMIL pooling module.

- **Cross-Lingual Multimodal Video Retrieval for Educational Content** *(Jul'22-Present)*
  *(Ongoing M.Tech Project in collaboration with **Google**, Guides: Prof. Preethi Jyothi, Prof. Ganesh Ramakrishnan)*
  - **Objective:** Given a query in English retrieve relevant videos in multiple languages.
  - We are working on exploiting speech and text modalities of technical content videos to generate a retrieval system using the content inside the video and not j ust the metadata.
  - Investigating **weakly-supervised multi-task** objectives to generate quality results without fine-grain annotations.

- **VaNI: Voice and Natural language models for Indian languages** *Ongoing *(Jan'23-Present)*
  (Guide: Prof. Preethi Jyothi)
  - Building a library for improved speech and language models for Indian languages.
  - Our goal is to make quality accented ASR and language models available without a hassle.
  - Exploring **Code-Switched ASR** in a **few-shot** setting.

COURSE PROJECTS

- **Hierarchical Table QA using LLMs***Ongoing *(Jan'23-May'23)*
  *(CS 728: Organization of Web Information, in collaboration with **IBM**, Instructor: Prof. Soumen Chakrabarti)*
  - Exploring Large Language Models (LLMs) like **Llama 7B** for answering complex questions with Hierarchical tables.
  - Complex questions require mathematical reasoning and hence, we are aiming to finetune/prefix-tune LLMs with noisy reasoning data.

- **SVM based Hate Speech Detection** *(Aug'21-Dec'21)*
  (CS 626: Speech, Natural Language Processing and the Web, Instructor: Prof. Pushpak Bhattacharyya)
  - Implemented binary classifier using **Support Vector Machines(SVM)** and its variants.
  - Performed **feature engineering**, explored **data augmentation**, and investigated **kernel approximations** for better speed.
  - Performed qualitative analysis that revealed the presence of **racial bias**, the effect of **negation**, immunity to **typographical error**, and the **effect of toxic words on non-toxic words.**
- **Earthquake Prediction** *(Aug'21-Dec'21)*
  (CS 725: Foundations of Machine Learning, Instructor: Prof. Preethi Jyothi)
  - Performed classification and regression tasks on **STanford EArthquake Dataset (STEAD).**
  - Studied the effect of feature engineering and encoding techniques like **hash encoding, cyclic features**, and hyperparameter tuning using **GridSearchCV** and **RandomSearchCV**.
  - Conducted analysis on algorithms like **Linear Regression (Lasso & Ridge), Decision Tree, XGBoost & Neural Network Regressor and Classifier.**
- **Few shot Image classification** *(Jan'22-May'22)*
  (CS 763: Computer Vision, Instructor: Prof. Sharat Chandran)
  - Developed architecture for **N-way K-shot** image classification.
  - Used **ResNet18** as feature extractor and techniques like **Autoencoder, Contrastive learning, and Self Attention.**
  - Analyzed the influence of N and K on the model performance.
- **N-Gram LM & ASR system** *(Jan'22-May'22)*
  (CS 753: Automatic Speech Recognition, Instructor: Prof. Preethi Jyothi)
  - Built N-Gram language model using **KenLM** and **OpenFST**.
  - Used language model to **fill the blank** given some prompt text; **generate a meaningful sentence** given some words in any order.
  - Developed Automatic Speech Recognition system for Wolof language using **Kaldi.**

TECHNICAL SKILLS

- **Programming & Scripting Languages**: C, C++, Python, SQL, HTML/CSS, JavaScript.
- **Tools & Libraries**: PyTorch, HuggingFace, NumPy, Pandas, Kaldi, LaTeX, Git.

POSITIONS OF RESPONSIBILITY

- **Academic Council Coordinator (Masters'), PostGraduate Academic Council (PGAC)** *(Sep'21-Apr'22)*
  - Coordinated with **40+** PG Class Representatives and PGAC members to resolve student queries.
  - Aided Executive members, General Secretary, and Institute Secretary of Academic Affairs (GSAA & ISAA) in conducting institute-wide events.
- **Interview Coordinator** *(Dec'21)*
  - Coordinated with a team of **250+** members for interviews of **1800+** students
  - Assisted in conducting Tests for **20+** firms and handling student queries.
- **CodeWars Mentor, Web and Coding Club (WnCC)** *(Dec'21)*
  - Mentored a group of **4** UG freshers to learn coding in Python and compete in CodeWars' bot programming contest.
  - Contributed a Python script with a primitive strategy accessible to all participants as starter code.
- **Teaching Assistantship**
  - **CS 753: Automatic Speech Recognition** *[*]Ongoing
    *(Instructor: **Prof. Preethi Jyothi**)* *(Jan'23-Present)*
    Contributing to the assessment of seminar-style paper presentations, and resolving student queries.

  - **CS 725: Foundations of Machine Learning**
    *(Instructor: **Prof. Preethi Jyothi**)* *(Jul'22-Dec'22)*
    Assisted the instructor with the evaluation, and solved student queries for a class of **300+** students

  - **CS 101: Computer Programming and Utilization (×2)**
    *(Instructors: **Prof. Parag Chaudhuri, Prof. S Akshay**)* *(Nov'21-Jun'22)*
    * Guided **26** students (14+12) spanning both iterations to learn to code in C++ language.
    * Tracked progress of the students and participated in help sessions to accelerate student performance.