

PARTH LAD

Data Scientist · TTS, ASR & Speech Datasets

+91 7990854853 · parthlad2403@gmail.com · parth-dev.in

linkedin.com/in/parth-lad-372564245 · github.com/ParthLad03 · Vadodara, India · Open to remote

SUMMARY

Data scientist working across the full speech pipeline for Indic languages: dataset construction, model fine-tuning and evaluation. Build production pipelines for corpus creation — segmentation, denoising, source separation, diarization, quality scoring — and evaluate ASR against human-annotated ground truth using WER, CER and DER. Experience fine-tuning multilingual TTS and building large-scale synthetic speech corpora on GPU.

EXPERIENCE

Data Scientist — Datasetsmith.com

Jun 2025 – Present

Full-time · Vadodara, Gujarat, India

- Build and maintain production speech data pipelines (segmentation, denoising, source separation, diarization, DNSMOS quality scoring) for TTS corpus construction; pipelines have processed thousands of hours of audio.
- Built a 15-language parallel synthetic speech corpus using zero-shot voice cloning on OmniVoice — roughly 100 hours per language, with matched emotion and delivery styles across the full set.
- Trained a custom XTTS v2 model for production-grade Gujarati TTS, covering native speaker-led data collection, preprocessing and fine-tuning.
- Evaluate ASR systems against human-annotated ground-truth transcripts — WER and CER for single-speaker corpora, DER for conversational and multi-speaker data — using an in-house annotation tool.
- Run GPU-accelerated inference on CUDA across pyannote, demucs and denoising stages, using batching and VRAM management to keep the GPU saturated on long-running jobs.
- Design HITL annotation workflows pairing automated first-pass processing with human reviewer validation, and built a WebRTC application capturing dual-speaker conversations in stereo with live turn-diarized transcripts.

Python Developer Intern — GradsScaler Technologies Pvt. Ltd.

Nov 2024 – May 2025

Vadodara, Gujarat, India

- Built a Gradio-based annotation tool for HITL audio dataset validation, supporting data quality workflows for speech model training.
- Developed a web application for studio-quality audio recording, compatible with professional audio interfaces, for clean TTS corpus collection.
- Built RAG chatbots and LLM API integrations in Python.

PROJECTS

Hindi TTS — XTTS v2 Fine-Tuning

Fine-tuned XTTS v2 on a curated 50-hour Hindi corpus, covering collection, preprocessing and speaker embedding normalisation, with GPU training on RunPod.

XTTS v2, Coqui TTS, PyTorch, RunPod

English → Devanagari Transliteration Engine

Rule-based NLP engine converting English text to Devanagari script at 86.6% word-level exact match. Uses a seven-tier lookup cascade running from a 7,300-word dictionary through to a neural phoneme model, handling silent letters, suffix morphology and multi-word phrases.

Python, NLP, Flask, REST API, neural phoneme model

Zero-Shot Voice Cloning Corpus

A 15-language parallel synthetic speech corpus built with zero-shot voice cloning on OmniVoice, around 100 hours per language. Each language carries the same matched set of emotion and delivery styles — happy, sad, angry, storytelling, news read and commentary — giving downstream models consistent expressive coverage across the set.

OmniVoice, zero-shot voice cloning, synthetic data, Python, RunPod

TECHNICAL SKILLS

Programming: Python, JavaScript

Speech & Audio: XTTS v2 architecture, LoRA adapters, base model fine-tuning, OmniVoice, Coqui TTS, Whisper, pyannote, demucs, FFmpeg, Pydub, Librosa, DNSMOS, speaker diarization, VAD

Evaluation & Data: WER / CER / DER, RTFx and latency profiling, MOS, MUSHRA, F1, BERTScore, speech corpus development, HITL annotation workflows, human-annotated ground truth

ML: PyTorch, NumPy, Pandas, YOLO v26, RAG pipelines and vector memory, LLM integration, FastAPI, Gradio

Infrastructure: CUDA, NVIDIA GPUs, Docker, Linux/Debian, virtual machines, RunPod, S3, Cloudflare R2, Git, GitHub

Web & Backend: Flask, Streamlit, React, Node.js, REST APIs, MongoDB, PostgreSQL

Languages: Gujarati (native), Hindi (professional), English (professional)

EDUCATION

B.Tech, Computer Science & Engineering — ITM SLS Baroda University · Lateral entry, 3rd year 2025 – 2028

Diploma, Computer Science & Engineering — ITM SLS Baroda University · CGPA 8.26 2022 – 2025