Sofi Altamsh Shah *Machine Learning Engineer*

✉ sofialtamsh123@gmail.com ☎ +91-9156704982 📍 Bengaluru, Karnataka 🔗 Portfolio

🔗 linkedin.com/in/sofialtamsh 🔗 github.com/sofialtamsh

Professional Summary

Machine Learning Engineer with 2+ years of experience building LLM and deep-learning models, now focusing on speech and audio ML using Whisper, torchaudio, and feature extraction. Skilled in fine-tuning models, optimizing inference, and developing scalable ASR/TTS pipelines. Proven impact by cutting QnA response time by 92% and improving hiring outcomes by 23% through production-ready ML systems.

Experience

Machine Learning Engineer, *Masai*   01/2025 – present | Bengaluru
- Built an automated **multilingual audio-generation** pipeline that converted recorded sessions into five languages using ASR, translation, and TTS models, reducing manual effort by over **80%**.
- Designed and built multi-stage **RAG + AI-agent pipelines** using LangChain for intent extraction, context retrieval, and automated support ticket responses.
- Built a RAG-based answering system for recorded sessions, reducing query time by **92% (≤3.2s)**. Improved student hiring rate by **23%** using a CTC prediction model. Trained and deployed to production.

Data Analyst Associate → Data Research Analyst, *Masai*   04/2024 – 01/2025 | Bengaluru
- Build data pipelines by analyzing the job market data from (LinkedIn, Naukri, Instahyre), delivering insights that improved placement rates by **15%**. Analyzed trends and built dashboards.

Data Science Intern, *1Stop.ai*   06/2022 – 11/2022 | Bengaluru
- Trained predictive models for house-loan value achieving more than **90% accuracy**, enhanced performance using **PCA**, and built insight dashboards with clustering visualizations.

Technical Skills

Audio & Speech ML: Whisper, torchaudio, librosa, MFCCs, VAD, augmentation, noise filtering

Machine Learning: Regression, Clustering, PCA, XGBoost, Feature Engineering, Hyperparameter Tuning

Deep Learning: CNN, RNN, GANs, Transformers (RoBERTa, T5, GPT), LoRA

NLP & LLMs: Prompt Engineering, RAG, Text2Text, NER, NLTK, Hugging Face, Fine-Tuning, LangChain

MLOps: Streamlit, Flask, HDFS, AWS Lambda, Vector Databases (FAISS), MySQL, MongoDB, Docker

Libraries & Frameworks: NumPy, Pandas, TensorFlow, PyTorch, Scikit-learn, SciPy, OpenCV, OCR

Visualization: Power BI, Matplotlib, Seaborn, Excel, Metabase

Projects

Multilingual Speech-to-Text Enhancement
- Fine-tuned Whisper on a small English–Hindi audio dataset using torchaudio for trimming, silence removal, MFCCs, and spectrogram generation.
- Improved WER by **22%** through targeted audio augmentation and built a fast real-time inference script optimized with fp16 and caching.

Fake News Detection Using Multi-Modal Learning
- Built a hybrid multimodal model (RoBERTa + CNN) for fake news detection, achieving **88.7% accuracy** on a 10K+ sample dataset using PyTorch & Hugging Face.

Education

B.Tech in Computer Science Engineering (8.1 CGPA),   2019 – 2023
Government College of Engineering, Aurangabad