

Promila Ghosh

✉ promilaghoshmonty@gmail.com

in promila-cse | pr0mila

🔗 https://promila.info/ | 📍 Dhaka, BD

EXPERIENCE

Golden Harvest Infotech

Senior AI Engineer

Dhaka, Bangladesh

Sept 2025 – Present

- Led the end-to-end architecture, development, and evaluation of AI-powered RCM development **Hppy Claim** and evaluation for the **USA, UAE, and India markets**, covering **75,000 claims and 82 model runs** with **0% cross-tenant leakage and 0% invalid strategy references**.
- Designed an optimized Denial Intelligence and Strategy Engine by calculating denial patterns in **SQL first**, reducing LLM calls from **15 to 1 per analysis** with **72-hour tenant-aware caching** and deterministic fallback, while sending only aggregated data to the LLM.
- Architected a multimodal healthcare document **HealthScan** intelligence platform, achieving **90% extraction accuracy, 90% agent task success, and 85% retrieval Hit@1** across mixed-format medical documents.
- Drove HealthScan AI to production readiness with **2% OCR CER and 40s p95 latency**, passing the **GO/NO-GO gate with zero safety leaks**.
- Designed a deterministic **fallback layer achieving 100% fallback success**, ensuring reliable strategy resolution.

United We Care

Data Scientist

July. 2022 – March, 2025

Gurgaon, Haryana, India (Remote)

- Designed and developed a transformer ASR system **United-MedASR** for medical transcription, reducing **Word Error Rate (WER)** to **5.96%** and improving transcription speed by **7-10x** with an optimized interfacing mechanism.
- Built a **5,486 hours** of synthetic clinical audio using a **Diffusion Model** and **GPT**, featuring male and female voices. Open-sourced as **United-Syn-Med** on **Hugging Face** for research.
- Developed a **multi-modal agent system** for **production** using **LangGraph** for analysing and classifying complex medical documents in various formats (*PDFs, scanned images, and handwritten notes*) using a **Large Multi-Modal Model (LMM)**, achieving **70% accuracy** on multi-page and handwritten inputs.
- Built systems for **face detection, gesture recognition, and facial analysis**, achieving around **60% accuracy** and **faster detection** with **sliding windows** and **Redis** for **real-time processing**.
- Streamed real-time user data into **BigQuery** for a **10K+ download app**, improving analytics and data management.

SKILLS

Deep Learning : Neural Networks, Transformers, Large Language Models, Reasoning Models, Vision-Language Models

LLM & ASR Technologies : LLaMA, Gemini, GPT, Whisper, Gemma, Retrieval-Augmented Generation (RAG), AI Agents

Tools & Frameworks : Python, Scikit-Learn, PyTorch, LangChain, LangGraph, FastAPI, Kafka, Groq, Fireworks AI, Nebius

Databases & Data : PostgreSQL, MongoDB, BigQuery, Redis, Amazon S3, Qdrant

Cloud Computing : AWS (EC2, S3), GCP (Compute Engine, Vertex AI, Google Cloud Storage)

PROJECTS

- **Medical Document Processing & ASR United We Care** Developed an intelligent system for ASR and medical document analysis using transformers.
- **MediBeng Whisper Tiny: Code-Switched Bengali-English Speech Translation for Clinical Settings** Built a dataset and fine-tuned the Whisper Tiny model for transcription and translation of code-switched speech in healthcare.

EDUCATION

M.Sc. Engg., Computer Science and Engineering

Khulna University – CGPA: 3.88/4.00 (**Distinction**)

February 2022 – August 2023

Khulna, BD

Publications

- **Recognition of Sunflower Diseases Using Hybrid Deep Learning and Its Explainability with AI – Mathematics (MDPI), 2023**
- **Risk prediction of ischemic heart disease using artificial neural network – IEEE, 2018**
- **Multi-labelled Bengali Public Comments Sentiment Analysis with Bidirectional Recurrent Neural Networks – Taylor & Francis, 2023**