




Promila Ghosh

✉ me@promila.info |  promila-cse |  pr0mila |  promila.info | 📍 Khulna, BD

EXPERIENCE

United We Care Data Scientist

July, 2022 – March, 2025
Gurgaon, Haryana, India

- Fine-tuned A **Transformer ASR system** for medical transcription, reducing **Word Error Rate (WER)** to **5.96%** and improving transcription speed by **7-10x** with an optimized interfacing mechanism.
- Generated **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.
- Designed and implemented a medical document analyzer and classifier using a **Large Multi-Modal model (LMM)**, achieving **around 100% accuracy** for multi-page and handwritten documents.
- Built **multi-modal agents** using **LangGraph** for **real-time image analysis** and **document transcription**, following an **event-driven architecture** to improve accuracy with categorized prompts and prepare outputs for **RAG** in a vector database.
- Developed systems using **DeepFace** and **MediaPipe** for **face detection**, **gesture recognition**, and **facial analysis**, achieving **97% accuracy** and **faster detection** with **sliding windows** and **Redis** for **real-time processing**
- Created a user emotion analysis system that integrates **text**, **speech**, and **video data** to assess emotional states, enabling **real-time control** of a **bots expressions** and **actions**.
- Built a robust data pipeline to seamlessly stream real-time user data into **BigQuery** for a high-traffic app (**10K+ downloads**), enhancing analytics efficiency and optimizing data management.

SKILLS

Deep Learning : Neural Networks, CNNs, RNNs, LSTM, Transformers (BERT, GPT), Named Entity Recognition (NER), Sentiment Analysis, Large Language Modeling, Chain of Thought, Reasoning Models

Multimodal Models & LLM : LLaMA Vision, LLaVA, Gemini, MediaPipe, Deepface, Deepseek, GPT

Speech Recognition (ASR) : Whisper, Wav2vec, Speech-to-Text APIs

APIs & Integrations : OpenAI API, Google API, GROQ API, FastAPI

Event-Driven Architecture : Kafka, Publish/Subscribe Model, Real-Time Stream Processing, Consumer Groups

Caching & Databases : MongoDB, PostgreSQL, Redis, Neo4j, BigQuery

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

Programming & DevOps: Python (Pandas, NumPy, Scikit-Learn, TensorFlow, PyTorch, Langchain), Docker, Git

PROJECTS

- **MediBeng Whisper Tiny: Code-Switched Bengali-English Speech Translation for Clinical Settings**
Built a dataset and fine-tuned Whisper Tiny model for transcription and translation of code-switched speech in healthcare.
- **Medical Document Processing United We Care | A SuperApp for Mental Wellness**
Developed intelligent agent for medical document classification using multimodal and image processing.

EDUCATION

M.Sc. Engg., Computer Science and Engineering Khulna University

February 2022 – August 2023
Khulna, BD

CGPA: 3.88/4.00, **Distinction**

Publications

- **Recognition of Sunflower Diseases Using Hybrid Deep Learning and Its Explainability with AI**
- **Risk prediction of ischemic heart disease using artificial neural network**
- **Multi-labelled Bengali Public Comments Sentiment Analysis with Bidirectional Recurrent Neural Networks**