

Emtiaz Uddin Ahmed

Machine Learning Engineer & Independent Researcher

Reliable LLMs & agents | VLM uncertainty | Low-resource speech

✉ emtiaz4546@gmail.com | 🌐 emtiaz-ahmed.pages.dev | 📍 Dhaka, Bangladesh

🔗 [Google Scholar](#) | 🔄 [EMTIAZ-RUET](#) | 📄 [emtiazuddinahmed](#) | 📦 [EMTIAZZ](#)



RESEARCH INTERESTS

I study how modern AI systems fail under real-world conditions, and how to detect and repair those failures without retraining. My recent work asks (i) when large language models and agents *state things they were never told*, by filling required structured-output fields with invented values and by recording their own inferences as user-stated facts in long-term memory; (ii) how much a vision-language model *knows about its own errors*, from anti-calibrated confidence in diffusion VLMs to mechanistic diagnoses that do not convert into fixes; and (iii) why multilingual text-to-speech models *fail on the long tail* of rare letters and newly added vocabulary. Building production voice-AI and agentic systems in industry is where these questions come from; I want to study them rigorously as a PhD student, starting Fall 2027.

Keywords: trustworthy LLMs and agents, abstention and hallucination, agent memory, uncertainty quantification and calibration, interpretability of VLMs, multilingual and low-resource TTS, evaluation methodology.

EDUCATION

B.Sc. in Computer Science and Engineering

Rajshahi University of Engineering & Technology (RUET)

2018 – 2022

Rajshahi, Bangladesh

MANUSCRIPTS

Five manuscripts on LLM and agent reliability, vision-language model uncertainty, and multilingual speech synthesis are currently under double-blind review. Titles and drafts are available on request.

RESEARCH EXPERIENCE

Independent Researcher

Self-directed, alongside full-time industry work

Sep 2025 – Present

Dhaka, Bangladesh

- Design and run pre-registered measurement studies of LLM and agent reliability across many open-weight and frontier models, using deterministic, judge-free metrics where possible.
- Study uncertainty and failure mechanisms in vision-language models with probing, activation patching and controlled interventions.
- Run controlled experiments on multilingual TTS, and release benchmarks, code and audio with each paper.

INDUSTRY EXPERIENCE

Senior Machine Learning Engineer

Tulip Tech Ltd.

Apr 2026 – Present

Dhaka, Bangladesh

- Led the migration from a monolithic to a microservices architecture, improving deployment frequency by 60%.
- Built *Ask Empathika*, an agentic RAG system using multi-agent orchestration for care assistance.
- Developed automated care-plan generation, risk-assessment and incident-reporting modules, reducing manual documentation by 50%; built a staff-prediction system with a custom fairness-scoring algorithm.

Senior Machine Learning Engineer

Markopolo AI Inc.

Sep 2025 – Mar 2026

Dhaka, Bangladesh

- Built a custom voice-AI pipeline on open-source TTS, replacing commercial providers and reducing per-call cost from \$0.13 to \$0.02; implemented voice-call orchestration with custom Bengali TTS.
- Architected a multi-channel content-generation engine (Email, SMS, WhatsApp, Voice, Push) with 30+ language localization, and a Graph-RAG retrieval layer.

- Built a Text-to-SQL assistant with automatic schema understanding, and a real-time license-plate detection system (fine-tuned YOLO + Kafka, 97% detection accuracy in production).
- Built a real-time AI interview system on LiveKit and a RAG-based conversational property search.

OPEN MODELS & DATA (HUGGING FACE)

Accounts: huggingface.co/EMTIAZZ (main) huggingface.co/EmtiazUdd (research data and checkpoints)			
Bangladeshi Bangla VITS TTS: Bangladeshi-accent VITS; 300+ downloads/month		model demo	2025
code			
Orpheus-3B Bangla: high- and low-data Bangla fine-tunes of an LLM-based TTS model		model	2026
Bangla XTTS-v2 voice cloning: interactive demo		demo	2026
Bangla 60-voice speech: multi-speaker Bangla speech datasets		full chunked	2026

HONORS & AWARDS

IUT 11th National ICT Fest, IUPC: ranked 18th of 150+ teams	2024
HackNSU: top-10 finalist of 100+ teams (real-time disaster-relief platform)	2023
ICPC Asia Dhaka Regional: ranked 33rd of 150+ teams	2022
ICPC Asia Dhaka Regional Preliminary: ranked 46th of 1,600+ teams	2022
National Hackathon on Frontier Technologies: finalist of 2,000+ teams (AI traffic monitoring)	2022
RUET CodeSmash: ranked 5th of 100 contestants	2019
Competitive programming: 3,000+ problems solved. LeetCode max 1982, top 7.3% (emtiazahmed) Codeforces Specialist, max 1525 (dakshg) CodeChef max 1731 (knight_01) AtCoder (EMTIAZ_AHMED)	

TECHNICAL SKILLS

Research	Experimental design and pre-registration, statistical testing, activation patching and probing, LoRA fine-tuning, uncertainty quantification, human and automatic evaluation of speech
ML	PyTorch, Hugging Face (Transformers, PEFT), JAX, vLLM, MLX; LLMs, VLMs, TTS/STT, RAG, multi-agent systems
Engineering	Python, C++, Java; FastAPI, LangGraph, LiveKit, Docker, Kubernetes, AWS, Kafka
Mathematics	Linear algebra, probability and statistics, optimization