

AL JAMI ISLAM ANIK

+8801790801923 | anik.islam1494@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

PROFILE

I am a software engineer who enjoys building simple solutions to complex problems. I have experience building production-grade software from scratch through to deployment, and I am particularly keen on the theoretical and implementation aspects of Small Language Models and RAG systems.

EDUCATION

BRAC University

Bachelor of Science in Computer Science – CGPA: 3.72 / 4.00

Dhaka, Bangladesh

2022 – 2026

PUBLICATION

AgentCollabBench: Diagnosing When Good Agents Make Bad Collaborators | [arXiv:2605.08647](#)

- *ICML 2025 Workshop on Failure Modes in Agentic AI*. Introduced a 900-task diagnostic benchmark spanning software engineering, DevOps, and data engineering to measure four behavioral failure modes in multi-agent LLM systems: instruction decay, false-belief contagion, context leakage, and tracer durability.

WORK EXPERIENCE

Software Engineer

Ask Turing AI

July 2024 – April 2026

Remote, United States

- Helped building a **RAG system** covering document ingestion, retrieval pipelines, and citation-grounded response flows.
- **Designed and implemented a complex file management system** with role-aware access, versioning, and bulk operations.
- Helped developing an **automated scheduled task system**, enabling non-technical users to configure, monitor, and manage background jobs.
- **Engineered an end-to-end authentication system**: email/password, **2FA**, **SSO**, and **RBAC** with fine-grained permission controls.

PROJECTS

Saul Goodman | *Python, Chroma, BM25, MinerU*

[GitHub](#)

- A grounded **RAG CLI** that answers questions claim-by-claim with per-chunk citations. Built a **hybrid retrieval pipeline** (dense Chroma + BM25 with **Reciprocal Rank Fusion** and **cross-encoder reranking**) and a **faithfulness auditor** that flags unsupported claims. Operator edits feed back as reusable style, domain, and citation-trust signals, so the system improves with use.

Forgotten Lily | *Python, Gradio, Qwen2.5-7B, Nemotron-Mini-4B, Modal*

[GitHub](#) | [Live](#)

- A narrative mystery game for the HuggingFace Build Small Hackathon. **Orchestrates three tiny models on a single T4 GPU** via Modal: **Qwen2.5-7B** selects tone IDs from a constrained vocabulary, **Nemotron-Mini-4B** grades player guesses, and **Kokoro TTS** renders a spoken line. Deterministic server-side game logic makes the system robust to hallucination.

Qwen2-VL Fine-tuning | *Transformers, Unsloth, QLoRA*

[Kaggle](#)

- Fine-tuned a **7B-parameter vision-language model** using **QLoRA** for memory-efficient training, achieving **80% accuracy** through systematic hyperparameter tuning and building the full data pipeline end-to-end.

Hydroharvest | *FastAPI, Scikit-learn, Random Forest*

[GitHub](#)

- Built for **NASA Space Apps Challenge 2024**. Trained a **Random Forest** regression model for rainfall prediction and served it via a **FastAPI** inference endpoint, covering the full ML workflow from feature engineering through deployment.

SLM From Scratch | *PyTorch, Python*

[GitHub](#)

- Implemented a **small language model** from scratch to understand **transformer internals**: tokenization, attention, positional encoding, and the training loop, without relying on high-level abstractions.

TECHNICAL SKILLS

AI / ML: RAG Systems, LangChain, LangGraph, LLM Fine-tuning (QLoRA, Unsloth), PyTorch

Data Stores: Chroma DB, Pinecone, BM25, PostgreSQL, Supabase

Languages: Python, TypeScript, JavaScript, SQL, Bash

Backend & Infra: FastAPI, Docker, GitHub Actions, CI/CD Pipelines

Frontend: React, Next.js, Tailwind CSS, Shadcn, Framer Motion

ACHIEVEMENTS

- 1st Place, BRAC University Intrachactive 1.0 Hackathon
- 5th Place, Innovation Cohort 2 (BUET Rise)