Al Jami Islam Anik

Dhaka, Bangladesh anik.islam1494@gmail.com Portfolio LinkedIn GitHub +8801790801923

Research Interests

I am broadly interested in the **reliability of language models**, particularly **small language models** and the **multi-agent systems** built from them. I am also interested in **multimodal understanding** and **low-resource languages** such as Bengali. I bring over two years of production software engineering experience, building and shipping LLM-powered systems end to end.

Education

BRAC University 2022 – 2026
B.Sc. in Computer Science and Engineering *CGPA: 3.72/4.00 (High Distinction)*
Thesis Supervisor: **Dr. Jannatun Noor Mukta**

Coursework: Machine Learning, Artificial Intelligence, Natural Language Processing, Linear Algebra, Statistics

English Proficiency: IELTS Academic, **Overall Band 8.0**, Nov. 2026

Research & Publications

Peer-Reviewed Publications

AgentCollabBench: Diagnosing When Good Agents Make Bad Collaborators. ***NeurIPS 2026***, Datasets and Evaluations Track; ***ICML 2026*** Workshop on Failure Modes in Agentic AI (FAGEN).
A. Mazumder, S. R. Dipta, N. J. Lia, …, S. A. Rani, **A. J. I. Anik**, A. N. Khan

- A 900-task benchmark for diagnosing how multi-agent LLM systems fail when agents share information.
- Designed the **cross-task privacy leakage evaluation**, which checks whether private context from one task leaks into a later task in the same session, and helped annotate the dataset.

Manuscripts Under Review

Analyzing the Semantic Roles in Bangladeshi Memes: A Multimodal Approach to Contextual Role Labeling and Explanation Generation. Under review at ***Engineering Applications of Artificial Intelligence***.
A. J. I. Anik, A. S. K. Mazlish, S. M. Nabi, M. Ahsan, J. Noor

- Led this work as first author at the **C2SG Research Group**. Built a dataset of Bengali memes in which each person or group is labeled as a hero, villain, or victim, and used multimodal models to identify these roles and explain them.

Work Experience

Casemax Plus Remote, United States
Software Engineer *May 2026 – Present*

- Working with lawyers to build a legal AI product for mortgage foreclosure teams.
- Conducting **user interviews with attorneys** to learn their existing workflows, then building interfaces around those workflows so attorneys trust the AI enough to use it in their foreclosure work.

Askturing.ai Remote, United States
Software Engineer *Jul. 2024 – Apr. 2026*

- Built the Askturing.ai web app from scratch, leading its UI/UX and database design, and helped build the **agentic workflows** behind the core product.
- Built a **retrieval-augmented generation system**, from document ingestion to answers grounded in citations.

Projects

SLM From Scratch | *PyTorch, Python*

- Implemented a small language model from scratch in PyTorch to understand how transformers work, from tokenization and attention to the training loop.

Forgotten Lily | *Gradio, Qwen2.5-7B, Nemotron-Mini-4B, Modal* | **Best Demo, HF Build Small Hackathon 2026**

- A detective game where players question Lily, who can only answer with 28 glyph "tones," and work out what the tones mean to recover her forgotten memory. Live demo.
- Runs **three small models on a single T4 GPU**. Qwen2.5-7B writes Lily's replies from her fixed vocabulary, Nemotron-Mini-4B screens player input and grades each guess, and Kokoro TTS gives Lily her voice.
- Limiting Lily to a fixed vocabulary and checking every input with a separate judge model keeps the small models on track and limits hallucinations.

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

- Fine-tuned Qwen2-VL (7B) with QLoRA to classify entity roles in memes.

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

- A RAG command-line tool that answers questions claim by claim with citations, using hybrid retrieval with reranking and an auditor that flags unsupported claims.

Technical Skills

Machine Learning: PyTorch, Hugging Face Transformers, Unsloth, LoRA / QLoRA fine-tuning, Scikit-learn
LLM Systems: RAG (hybrid retrieval, BM25, reranking), LangChain, LangGraph, Gradio, Modal
Programming Languages: Python, TypeScript, JavaScript, SQL, Bash
Data & Infrastructure: PostgreSQL, Supabase, Chroma DB, Pinecone, FastAPI, Docker, CI/CD
Web Development: React, Next.js

Honors & Awards

- **Best Demo Award** ($1,000), HF & Gradio Build Small Hackathon 2026 (global), for Forgotten Lily. Official results
- **Vice-Chancellor's List** (4 semesters) and **Dean's List** (3 semesters), BRAC University.
- **1st Place**, BRAC University IntraHacktive 1.0 Hackathon.
- **5th Place**, Innovation Cohort, BUET Rise, where we also raised initial funding for our VR product.

References

Dr. Jannatun Noor Mukta
Associate Professor, Department of CSE
Director, BSc in Data Science Program
United International University, Dhaka, Bangladesh
jannatun@cse.uiu.ac.bd
Undergraduate thesis supervisor and co-author

David Veillon
Chief Executive Officer, Casemax Plus
Founder & CEO, Idea Builders LLC
Longmont, Colorado, United States
david@ideabuilders.studio
Current employer (CEO, Casemax Plus)