Aritra Mazumder

Salt Lake City, UT | aritra741@gmail.com | Portfolio | Linkedin | Github | +1-385-235-9286

Education

University of Utah **Aug '25 – Present**

Ph.D. in Computer Science – Advisor: Dr. Anna Fariha

- Research Focus: Automated relational database generation from unstructured text

Research Publications and Papers Under Review

- **QuWARTS: Query Workload Aware Relational Table Synthesis from Unstructured Text**. Accepted at *VLDB 2026*.
- **AtomSQL: Interactive Disambiguation of NL-to-SQL via User-Guided Atom-Level Alignment**. Accepted at *VLDB 2026*.
- **AgentCollabBench: Diagnosing When Good Agents Make Bad Collaborators**. Under review at *a double-blind top AI conference*. Pre-print available:
- **Exploring Optimal Placement of Head-based Hierarchical Marking Menus on Smartphones**. Published at *EAI MobiQuitous 2023*.
- **garNER at SemEval-2023: Simplified Knowledge Augmentation for Multilingual Complex Named Entity Recognition**. Published at *SemEval@ACL 2023*.

Work Experience

Askturing.ai - *Remote, California* **Feb '25 – May '25**

Senior Software Engineer

- Architected and deployed production **NL2SQL RAG pipeline** using **LangChain** and **LangGraph**, enabling natural language database queries for enterprise clients.
- Engineered legal-tech case resolution system with **Neo4j knowledge graphs**, processing complex relationship analysis across 100K+ legal documents.
- Optimized RAG retrieval pipeline latency by **60%** through parallel computation and semantic chunking improvements.
- Implemented multi-tool **AI agents** with capabilities for web search, code execution, data processing, and vector-based retrieval.

Blue Dot Intel - *Remote, New York* **Oct '23 – Feb '25**

Founding Software Engineer

- Developed **Retrieval Augmented Generation (RAG)** pipeline for investment data analysis, utilizing **Pinecone** vector database with 500K+ embeddings for efficient semantic search.
- Built full-stack architecture with **Next.js**, **MySQL** for relational data, and **Tailwind CSS**.
- Deployed scalable system on **AWS EC2** with **Nginx** reverse proxy and **Docker** containerization,
- Optimized the metadata tagging system performance by **13x** using **Aho-Corasick** algorithm.

Re:tune - *Remote, Delaware* **Sep '23 – Oct '23**

Software Engineer (Short-term Contract)

- Debugged and optimized production **RAG platform** built with **Next.js**, handling **100K+** daily queries.
- Reduced critical production bugs by **20%** and streamlined support ticketing system, decreasing average resolution time by **70%**.
- Created comprehensive technical documentation for RAG workflows and feature implementations.

Daricomma - *Bangladesh* **Mar '23 – Sep '23**

Founding Software Engineer

- Designed scalable backend using **Hasura GraphQL** and **PostgreSQL**, supporting **50,000+ active users** and **200,000+ questions** with **sub-100ms** query response times.
- Built full-stack application with **Next.js** frontend and **NestJS** backend services.
- Developed authentication system with JWT, integrated payment processing, and implemented a referral system.

Somadhan - *Remote, Bangladesh* **Oct '19 – Oct '21**

Software Engineer (Part-time)

- Developed cross-platform mobile application using **Flutter**, deployed to 5,000+ users on Android/iOS.
- Built RESTful backend API with **Node.js** and **MongoDB**, handling **50K+** daily requests with Redis caching layer.
- Created responsive admin dashboard with **React.js** for real-time analytics and user management.

Achievements

- Ranked **Candidate Master** on Codeforces (max rating **1932**, **Top 0.2%** globally out of **1.3M+** users, **Profile**).
- **ICPC Asia West Finalist 2021** – **5th place** (top 0.2%) at ICPC Dhaka Regional 2021, **10th place** (top 0.4%) at ICPC Dhaka Regional 2022 out of **2500+** teams.
- **Meta Hacker Cup 2022**: Qualified for Round 3 as **371st globally**, **top 1%** out of **27,500+** registrants. **Scoreboard**
- **6-star** on CodeChef (max rating 2244) – Top 10 nationally, **top 0.01%** globally out of **1M+** users. **Profile**
- **Top 1%** on LeetCode (max rating 2191) – Top 10 nationally. **Profile**
- **Problem Setter & Judge**: **HackerEarth** and multiple inter-university programming competitions

Projects

FindResearch.online | *Next.js, TypeScript, DistilBERT* | ***1000+** monthly users across **20+** countries*
- Research paper aggregation platform with **AI-powered insights** using **DistilBERT** transformers for automatic paper summarization.
- Implemented **BM25 relevance ranking** and advanced filtering.

Qualytics | *TypeScript, VS Code Extension API*
- VS Code extension for code quality visualization, analyzing **Cyclomatic Complexity**, **Maintainability Index**, and other code quality metrics.
- Features interactive Chart.js visualizations and project-level metrics aggregation for TypeScript codebases.

Dynamic Programming Simulator | *JavaFX, JFoenix*
- Educational desktop application visualizing 6 classic DP algorithms (**0-1 Knapsack, Edit Distance, Matrix Chain Multiplication and more**) with step-by-step animation.
- Implemented real-time speed control, state visualization, and pause/play control.

SonicAI | *Flask, TensorFlow, React.js*
- Full-stack music classification web app using **Convolutional Neural Networks** for genre, note, and emotion recognition.
- Averaged 90% accuracy on GTZAN dataset through spectral feature extraction and custom CNN architecture.

Leadership, Community Impact & Recognition

Hivemind Research Organization
Founder & Technical Lead — *May '25 – Present*
- Founded and lead research organization of 50+ PhD students, professors, and software engineers conducting collaborative research in RAG systems, agentic AI, data management, and foundation models.
- Coordinate weekly meetings, assign research tasks, and oversee multiple ongoing research projects across institutions including **UMBC**, **ASU**, and **UTD**.

Technical Education & Mentorship
Community Leader – 1250+ Members — *2024 – Present*
- Built and lead technical community of **1250+ competitive programmers and software engineers**, where I answer the community's technical questions.
- Invited as **industry expert** by **World Bank-funded** initiative by **Bangladesh ICT Ministry** to train undergraduates in technical skills at 2 universities.

Author: Competitive Programming Textbook
Manuscript Under Review by Established Academic Publisher — *2025*
- Completed a full draft of a **comprehensive competitive programming textbook** covering over **1000 problems** with detailed explanations.
- Currently under editorial and expert review by a major Bangladeshi publisher (non-commercial academic collaboration).
- Received advance interest and recommendations from **university professors** who plan to adopt the book once published.

SUST Competitive Programming Training Camp
Coordinator and Trainer — *2021 – 2023*
- Coordinated and trained **500+** students in advanced algorithms and competitive programming techniques.

Technical Skills

Languages: Python, C++, JavaScript/TypeScript, Java, SQL, C
AI/ML: LangChain, LangGraph, TensorFlow, Pinecone, Neo4j, OpenAI API, Hugging Face
Web: Next.js, React.js, Node.js, NestJS, Flask, Express.js, GraphQL (Hasura)
Databases: PostgreSQL, MySQL, MongoDB, Redis, Vector Databases (Pinecone, Weaviate)
DevOps and Tools: Docker, AWS (EC2, S3), Git, Nginx, Linux, REST APIs
Mobile: Flutter/Dart