

TOPU KUMAR MONDOL

topukumar538@gmail.com | +880 1311-280754 | LinkedIn: topu-kumar-mondol | GitHub: topukumar538

EDUCATION

Rajshahi University of Engineering & Technology (RUET)

B.Sc. in Electrical & Computer Engineering (ECE)

Rajshahi, Bangladesh

Expected Graduation: Oct 2028

- Relevant Coursework: Data Structures & Algorithms, Object-Oriented Programming, Database Systems

PROJECTS

OpsIQ — AI-Powered Ops Intelligence Platform

GitHub

Python · FastAPI · LangGraph · LangChain · FAISS · PostgreSQL (async SQLAlchemy)

2026

- Verified root-cause classification accuracy by replaying 3 documented production incidents (GitLab 2017, Cloudflare 2019, AWS 2020) through the pipeline, asserting both category matches and domain-specific terms drawn from the published postmortems
- Built a 5-node LangGraph DAG where log analysis and timeline extraction run in parallel and join at root-cause inference, over three temperature-tuned LLM instances (0.7 chat / 0.3 RAG / 0.1 postmortem) isolating generative, grounded and deterministic tasks
- Replaced per-request session lookups with stateless HMAC-SHA256 tokens carrying a version field, so bumping one database column invalidates every issued token without the server storing any of them
- Made restarts transparent by persisting FAISS stores to disk under per-user, per-session paths and mirroring conversation state to PostgreSQL, restoring summary buffers and recent messages on reconnect
- Automated the release path with a GitHub Actions pipeline that runs the 123-test suite on every push and deploys only when tests pass, replacing manual redeploys

ContentPlatform — Personalized Recommendation Engine

GitHub

FastAPI · PostgreSQL · SQLAlchemy 2.0 · JWT + OTP Auth · APScheduler · Docker · Pytest · Locust

2025

- Sustained a 99% success rate at 100 concurrent users (~800ms p95) by load testing with Locust from 10 to 500 users, and root-caused the failure at 500 as database connection pool exhaustion
- Stopped any one category from dominating the feed by capping it at 3 of the top 5 slots per page, using a slot-based ranker with softmax sampling at temp=0.7 and 10% random injection
- Kept recommendations following current behaviour rather than old activity by decaying interaction weights (view=1, like=3, save=5) on a 30-day constant, recomputed for every user in hourly batches
- Delivered a full 20 posts per page with zero repeats across the entire feed by building all pages in one pass against a shared seen-set, seeded so pagination stays stable
- Cut admin block latency from up to 7 days to a single request by re-checking user state in the database on every authenticated request instead of trusting the JWT
- Automated deployments to Hugging Face Spaces via a 58-test GitHub Actions CI/CD pipeline, eliminating manual releases

TECHNICAL SKILLS

Languages: Python, C++, C, SQL

AI / ML: LangChain, LangGraph, RAG, Retrieval Systems, Vector Databases (FAISS), Embeddings, LLM Integration

Backend & Databases: FastAPI, SQLAlchemy 2.0, PostgreSQL, REST APIs, JWT Auth, APScheduler

Cloud & Infrastructure: Docker, CI/CD (GitHub Actions), Hugging Face Spaces, Neon (managed PostgreSQL), Containerized Deployment

Tools: Git, GitHub, Pytest, Locust, Pydantic

CERTIFICATIONS

Machine Learning Specialization

Dec 2025

DeepLearning.AI / Andrew Ng (Coursera) · Credential ID: PTP0CHLR5O6S

CS50x: Introduction to Computer Science

Oct 2024

HarvardX (edX)

CS50P: Introduction to Programming with Python

2024

HarvardX (edX)

ADDITIONAL

- Solved 400+ data structures and algorithms problems across LeetCode, GeeksforGeeks and Codeforces — spanning arrays, trees, graphs, binary search and dynamic programming