

Anshuman Kumar

New York, NY | anshuman.k2021@gmail.com

linkedin.com/in/anshumankumarus | github.com/anshumankumar2021 | [Portfolio](#) | [LeetCode](#)

PROFESSIONAL SUMMARY

Software engineer and CS M.S. student with an Information Security background who builds secure, cloud-native backend services and LLM-powered applications. Hands-on with Docker, Kubernetes, CI/CD, AWS, and GCP, shipping low-latency search and chatbot systems, automated testing tools, and real-time Kafka transaction pipelines across the full software development lifecycle. Fluent with generative AI developer tools (Claude Code, LangChain).

TECHNICAL SKILLS

Languages: Python, SQL, Assembly, PySpark

Streaming, Cloud & DevOps: Apache Kafka, Docker, Kubernetes (GKE), AWS (Lambda, S3, DynamoDB, SQS, CloudWatch), GCP, Azure, CI/CD

Security & Observability: Secure coding, IAM/RBAC, least privilege, encryption at rest, prompt-injection filtering, Prometheus, Grafana

AI/ML & Data: LLMs (BERT, Whisper), RAG, LangGraph, LangChain, FAISS, Pinecone, Hugging Face, Spark ML, Delta Lake, MLflow

EDUCATION

The City College of New York, CUNY

Aug 2025 - Present

Master of Science in Computer Science | GPA: 3.88 | New York, NY

Vellore Institute of Technology

Sep 2021 - Jun 2025

B.Tech in Computer Science, Specialization in Information Security | CGPA: 8.91/10 | India

EXPERIENCE

Adjunct Lecturer

Jan 2026 - Present

Department of Computer Science, The City College of New York (CCNY) | New York, NY

- Teach undergraduate Assembly Language and Python to **120+ students** through lectures, hands-on projects, and labs.
- Design assignments, labs, and exams, and mentor students in debugging, algorithms, and industry-aligned coding practices.

Software Engineer Intern

Sep 2023 - Nov 2023

Tata Consultancy Services (TCS) | Gurugram, India

- Refactored front-end components, improving usability and perceived load times **~15%** across major browsers.
- Standardized reusable UI components, cutting production issues **~20%** while shipping features within a cross-functional team.

PROJECTS

Real-Time Transaction Fraud Monitor | [Live demo](#) | [GitHub](#)

Kafka, Python, Kubernetes, Grafana

- Caught **96.9%** of injected fraud incidents at **92.6%** precision across a 206K-transaction stream by building a **Kafka** pipeline that flags velocity bursts, amount spikes and impossible travel.
- Sustained **~49K msgs/s** at **61 ms p95** by scaling detectors as a consumer group on **Kubernetes** with Grafana alerts.
- Made retries safe via at-least-once delivery and a dead-letter topic, verified by end-to-end **Kafka** tests in CI.

Knowledge Base Assistant (RAG) | [Live demo](#) | [GitHub](#)

Python, BM25, embeddings, Vercel

- Ranked the right source first for **76.8%** of 10,570 held-out natural-language questions by building a hybrid **BM25 + dense-embedding** retriever with a learned re-ranker over 19K Wikipedia passages.
- Declined **83%** of out-of-scope questions by training a confidence model on retrieval scores, served at **13 ms p95**.

Agentic AI Workflow Orchestrator | [Live demo](#) | [GitHub](#)

LangGraph, Python, LLMs, SQL

- Answered **91.9%** of payment-risk questions correctly on a SQL-verified benchmark by building a **LangGraph** agent that plans, calls read-only SQL and policy-search tools, and verifies its answer.
- Allowed **zero** unauthorized actions under prompt-injection attacks by adding input screening and role-based tool gates.

Intelligent Document Processing Pipeline | [Live demo](#) | [GitHub](#)

Python, OCR, AWS Lambda, S3, DynamoDB

- Read receipt totals exactly right on **85.5%** of 200 unseen scans (**+7 pts**) via multi-pass **OCR** voting and learned scorers.
- Cut wrong values stored by **90%** while auto-processing **55%** of fields by routing low-confidence fields to human review, deployed as a serverless **AWS** stack with least-privilege IAM.

Customer Churn Lakehouse | [Dashboard](#) | [GitHub](#)

PySpark, Delta Lake, Spark ML, MLflow

- Made lookups **5.4×** faster and daily loads **46%** cheaper than rebuilds with a medallion **Delta Lake**, MERGE and Z-ordering.
- Passed data-quality gates on **36/36** daily batches with **0** duplicates on replay via row reconciliation and idempotent MERGE.
- Beat a recency-rule baseline by **6.5 AUC points** with a **Spark ML** model on point-in-time features, tracked in **MLflow**.

LEADERSHIP & CERTIFICATIONS

Leadership: CCNY Cybersecurity Club, Core Team (Media Management): run the club's media and moderate its Discord server.

Certifications: Linux Foundation: Kubernetes (2026); Anthropic: Claude Code (2026); Google Cloud: Generative AI (2026); Google Cloud Practitioner (2023); Hugging Face: AI Agents Course (2026); DeepLearning.AI: LangChain (2026); Python DSA + LeetCode (2025).