

Shutonu Mitra

mshutonu@vt.edu | 9296827538 | LinkedIn: shutonu-mitra | GitHub: sm5190 | sm5190.github.io

Education

Virginia Polytechnic Institute & State University

Master of Science in Computer Science

(Aug 2023 – May 2025)

GPA: 3.87

Military Institute of Science & Technology (MIST)

Bachelor of Science in Computer Science

(Feb 2018 – May 2022)

CGPA: 3.65

Work Experience

Back End Engineer, Curanostics Inc.

(May 2024–Aug 2024 | Atlanta, Georgia, USA)

- Designed and scaled patient-facing REST APIs as containerized microservices on **Google Cloud (Next.js, TypeScript, Docker, GKE)** behind an **API Gateway** and **load balancer**. Peak throughput **140** requests per second; p95 latency reduced from **360** ms to **160** ms.
- Delivered clinician dashboards in **React 18** with **React Query** and **Recharts**. Added optimistic updates, skeleton states, and error boundaries. Post-launch engagement up **25%**. Time to insight reduced **38%**.
- Improved developer velocity by establishing **GitLab CI/CD** pipelines and **Cloud Monitoring**, resulting in a **6x** weekly deployment frequency and an **18 – minute** MTTR.
- Productionized a medical report summarization and Q&A service using **Retrieval-Augmented Generation (RAG)** with **LangChain-based orchestration**, integrating prompt versioning, latency monitoring, and failure handling; improved medical text processing accuracy by **39%**.

Software Engineer Intern, Teletalk Bangladesh LTD.

(February 2021–April 2021 | Dhaka, Bangladesh)

- Engineered backend services for a nationwide exam-results platform serving **2M+** students on results day, running on **on-prem Linux VMs** behind **Nginx/HAProxy** with **MySQL**, ensuring reliable, low-latency access during peak traffic.
- Built the search feature in **Java (Spring Boot)** with composite indexing, pagination, and **Redis** caching; improved data retrieval speed by **26%**, enabling fast lookups under peak load.
- Implemented ingestion and validation services (**Java, Python/Node**) ; reduced mismatches by **72%** and duplicate inserts by **91%**, resulting in throughput increase from **22k** to **51k** records/min).
- Diagnosed and resolved **15+** critical defects across **Java** and **Python** components (race conditions, cache invalidation, null-safety, edge-case input handling); helped maintain **98.99%** uptime during release windows.

Data Engineer, Selise Digital Platform

(December 2021–March 2023 | Dhaka, Bangladesh)

- Designed and tuned **SQL (MySQL, Postgres)** and **NoSQL (DynamoDB)** schemas with composite indexing and connection pooling; cut p95 read latency from **140** ms to **90** ms and reduced database CPU by **20%**.
- Automated **Spark** and **Airflow** pipelines on **AWS (S3, Athena, EMR)** moving about **1 TB/day**; cut end-to-end runtimes by **45%** and reduced pipeline incidents by **40%**, with roughly **99%** SLA adherence.
- Containerized ETL components with **Docker** and deployed **Airflow** on **Kubernetes** with health checks, autoscaling, and rollbacks; cut deploy time from **20** to **8** minutes.
- Co-owned production model delivery with **Docker** and **Kubernetes**, maintaining approximately **220** ms p95 inference latency; shipped BI assets used by five teams, reducing time to insight by **42%**.

Graduate Research Assistant, Virginia Tech

(Aug 2023–May 2025 | Blacksburg, Virginia, USA)

- Built production-grade **data pipelines and ML services (Python, PyTorch, SQL)** for large-scale cyber-fraud analysis; improved forecasting accuracy **25%** vs baseline with reproducible training and evaluation.
- Designed and shipped reliability-focused systems including an **RL agent** with success and retry metrics, catching **3 – 5** regressions pre-release, and a privacy-preserving sensing pipeline reducing false alarms **30%**.
- Productionized ML workflows on **AWS SageMaker** (training, registry, batch and real-time inference) with **p95** latency < **300ms** and automated retraining on data drift.

Projects

Cloud-native E-commerce App | React, Restful APIs, MySQL, JDBC, Tomcat, AWS

Github

- Owned a multi-tier e-commerce service with React SPA, Java/Tomcat backend with MySQL, deployed on AWS with VPC isolation and secure public/private subnet architecture. Optimized backend queries and connection pooling, reducing page load time by 30% and database response latency by 25% under concurrent traffic.

Serverless REST API with AWS Lambda | JAVA, AWS RDS, AWS Lambda, Terraform

Github

- Architected a serverless API layer using Lambda, API Gateway, and RDS Proxy to handle burst traffic with low operational overhead and fault isolation. Strengthened reliability with retries, DLQs, and automated IaC deployments, cutting average query latency by 30% and improving resilience during peak load.

Technical Skills

Coursework:

Data Structures and Algorithms, Operating Systems, Computer Networks, Distributed Systems

Languages:

Java, Python, Go, TypeScript, JavaScript, SQL, Bash, C, C++

Backend & APIs:

RESTful APIs, Microservices, Spring Boot, Node.js, Express, Django, FastAPI

Cloud & Infra:

AWS (EC2, S3, EMR, Lambda, RDS, CloudWatch, SNS, SQS), Google Cloud (GKE), Docker, Kubernetes, Terraform, GitHub Actions, Linux

Data & Messaging:

MySQL, PostgreSQL, DynamoDB, Redis, Spark, Airflow