

Nikhil Murali

Buffalo, NY

✉ nmurali3@buffalo.edu

📁 [Portfolio](#)

in [LinkedIn](#)

🐙 [GitHub](#)

☎ (716) 228-0446

Summary

Software Engineer with 3+ years of experience in full-stack development, distributed systems, and applied AI. Strong background in designing scalable platforms, optimizing system performance, and leading projects through the full SDLC.

Education

University at Buffalo, State University of New York

Masters of Science, in Computer Science (GPA: 3.75)

Aug 2024 – Dec 2025

Buffalo, NY

National Institute of Technology Karnataka, Surathkal

Bachelor of Technology, in Electrical and Electronics Engineering

July 2017 – April 2021

Surathkal, India

Professional Experience

Empire State Realty Trust

Software Engineer Intern, AI

June 2025 – Present

Manhattan, NY

- Developed an AI Assistant using Salesforce Agentforce to automate data retrieval, CRM updates, and tour scheduling for prospective tenants, streamlining the leasing workflow and saving brokers 20% of their time.
- Collaborated cross-functionally with Leasing stakeholders to identify and document pain points in Salesforce CRM usage.

Standard Chartered Bank

Software Development Engineer

July 2021 – July 2024

Bangalore, India

- Reduced authentication latency by 38% (from 2100ms to 1300ms) by decoupling [ThreatMetrix](#) into a standalone Java Spring microservice and implementing distributed caching with Hazelcast.
- Upgraded Java(11→18), restructured PostgreSQL schema to improve query performance by 25%, and integrated Kafka to enhance message throughput, enabling reliable processing of up to 1M events/day.
- Led the end-to-end SDLC of Kong API Gateway routes, leveraging Agile methodologies and cross-functional collaboration to onboard global financial clients into the Open Banking ecosystem.
- Deployed containerized API gateway application on AWS EC2 with Docker, Ansible and Azure DevOps, automating CI/CD pipelines for internal API routing. Created Python scripts for REST API testing, log/file parsing, and system health checks.
- Used ELK stack to monitor real-time API Gateway traffic logs, debug errors, and analyze usage patterns.

Key Projects

Ultra-Low-Latency Order Matching Engine (C++20) [🔗](#)

- Designed and implemented an ultra-low-latency C++20 order matching engine with price-time priority order book, supporting limit/market orders and deterministic replay via binary write-ahead logging.
- Optimized memory access to achieve sub-microsecond latency (150ns inserts, 250ns matches) and >100k ops/s throughput.
- Built robust infrastructure for testing and benchmarking, including GoogleTest unit coverage, Google Benchmark microprofiling, and crash-replay testing to ensure correctness under load.

Personalized Conversational AI for Enhanced AAC Communication [🔗](#)

- Developed a personalized conversational AI system for AAC users by fine-tuning FLAN-T5 and integrating RAG with user-specific narrative embeddings, enabling context-aware and empathetic responses.
- Implemented dynamic prompting and FAISS-based retrieval, with final response generation via Mistral LLM, resulting in a 16% improvement in BLEU and ROUGE scores. Built an interactive React frontend with Flask backend.

Toxic Audio Detection using LLMs (SUNY Research Foundation) [🔗](#)

- Worked on the development of a real-time voice chat moderation system for children's online games using multimodal LLMs. Designed a chain-of-thought (CoT) prompting architecture to improve toxic speech detection accuracy.
- Conducted a comparative evaluation of GPT-4o + CoT prompting against commercial moderation systems (Roblox, Perspective API), achieving a 30% higher F1-score (0.75) on a curated toxic/non-toxic audio dataset.

Technical Skills

Languages: Java, Python, C++, JavaScript, SQL, Bash, HTML, CSS

Web Development: Spring Boot, Node.js, JUnit, React, StreamLit, Flask, Docker, Kong, FastAPI

Cloud & DevOps: AWS, Azure DevOps, Kubernetes, Jenkins, Ansible

Data & Analytics: Kafka, PostgreSQL, MongoDB, Spark, Elasticsearch, Power BI, Tableau, MS Excel

AI/ML: PyTorch, TensorFlow, LangChain, MS Fabric, Agentforce, Keras, Scikit-learn, Numpy, Pandas

Certifications

- AWS Certified Solutions Architect - Associate [🔗](#) , AWS Certified Developer - Associate [🔗](#)
- Hugging Face - Fundamentals of Agents [🔗](#)