Nikhil Murali

New York, NY

📞 (716) 228-0446   ✉ nmurali3@buffalo.edu   💼 Portfolio   in LinkedIn   ⭘ GitHub

Summary

Full-Stack Developer with 3+ years of experience in backend systems, cloud infrastructure, and applied AI, currently pursuing an M.S. in Computer Science. Proficient in Java, Python, and JavaScript, with a proven track record of building scalable web and GenAI applications and leading end-to-end SDLC processes. Actively seeking SDE/SWE roles.

Education

University at Buffalo, State University of New York   Aug 2024 – Dec 2025
Masters of Science, in Computer Science (GPA: 3.75)   *Buffalo, NY*

National Institute of Technology Karnataka, Surathkal   July 2017 – April 2021
Bachelor of Technology, in Electrical and Electronics Engineering   *Surathkal, India*

Professional Experience

Empire State Realty Trust   June 2025 – Present
AI Software Engineer Intern   *Manhattan, NY*
- Building **AI agents** through **Azure AI Foundry** and **Salesforce Agentforce** to automate tenant leasing workflow.

Standard Chartered Bank   July 2021 – July 2024
Software Development Engineer   *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.
- Architected and deployed **containerized API gateway application** on **AWS EC2** with **Docker** and **Azure DevOps**, automating **CI/CD** pipelines for internal API routing.
- Created **Python** scripts for **REST API** testing (using Postman), log/file parsing, and system health checks. Also created **Ansible** playbooks for automated deployments.
- Used **ELK stack** to monitor real-time API Gateway traffic logs, debug errors, and analyze usage patterns.

Key Projects

Toxic Audio Detection using LLMs *(SUNY Research Foundation)* ⬈
- Spearheaded 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.

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.

NoteWiz: An AI-Powered Study Assistant ⬈
- Developed a **Streamlit-based** web application integrating **Mistral's LLM** for real-time study material generation (summaries, flashcards, quizzes). Optimized **few-shot learning** prompts to enhance the accuracy and relevancy.

Technical Skills

Languages: Java, Python, JavaScript, C++, SQL, Bash, HTML, CSS
Web Development: Spring Boot, Node.js, JUnit, React, Flask, Docker, Kong, PostgreSQL, MongoDB, FastAPI, Postman
Cloud & DevOps: AWS, Azure DevOps, Kubernetes, Jenkins, Ansible
Data & Analytics: Kafka, StreamLit, Spark, Elasticsearch, Power BI, Tableau, MS Excel
AI/ML: PyTorch, TensorFlow, Hugging Face, LangChain, Azure AI, Keras, Scikit-learn, Numpy, Pandas

Certifications

- **AWS Certified: Solutions Architect Associate** ⬈ , **AWS Certified: Developer Associate** ⬈