Bhavani Shankar Ajith

Stony Brook, NY
+1 (934) 949-8732 · abhavanishankar2002@gmail.com · github.com/abs768 · linkedin.com/in/abs768

Education

M.S. in Data Science *Expected: May 2026*
Stony Brook University (SUNY), Stony Brook, NY GPA: 3.71 / 4.00
B.Tech in Computer Science (AI and ML) *Graduated: June 2024*
SRM Institute of Science and Technology, Chennai, India GPA: 3.90 / 4.00

Skills

Programming Languages: Java, Python, SQL, JavaScript, C++, CUDA
Backend & Web: Spring Boot, REST APIs, Flask, Node.js, microservices architecture
Databases: PostgreSQL, MySQL; schema design, transactions, query optimization, indexing
Tools & DevOps: Git, Docker, Linux, JUnit, PyTest, CI/CD workflows
Cloud & Other: AWS, Azure, system design fundamentals
Performance Optimization: TensorRT, GPU profiling (Nsight Systems), hardware acceleration, memory optimization

Experience

Software Engineer Intern *June 2025 -- August 2025*
Navitas Business Consulting, Virginia *Spring Boot, PostgreSQL, JUnit*

- Developed and deployed Spring Boot REST APIs for member eligibility validation and enrollment workflows **supporting 184,000+ active members** across multiple health plans.
- Built PostgreSQL triggers and constraints to **automatically track enrollment changes and document updates, eliminating manual data reconciliation** and improving audit compliance.
- Implemented schema migration strategy using Flyway with **version control and rollback mechanisms, enabling safe database deployments** across development, staging, and production environments.
- Wrote comprehensive unit and integration tests using JUnit and Testcontainers to **validate business logic and database interactions**; contributed to code reviews and improved codebase maintainability.

Edge AI Systems Developer *May 2022 -- October 2022*
NVIDIA Jetson AI Industry Training, Chennai, India *Python, CUDA, TensorRT, Docker*

- Optimized vehicle damage detection CNN for deployment on NVIDIA Jetson edge devices; **achieved 87% accuracy with sub-100ms inference latency** through TensorRT optimization and FP16 quantization.
- Built hardware-accelerated preprocessing pipeline using VPI (Vision Programming Interface) to offload image operations from CPU to GPU, **reducing preprocessing bottleneck by 60%** and improving overall throughput.
- Profiled and optimized inference performance using Nsight Systems; **identified memory transfer bottlenecks and migrated to Unified Memory architecture to eliminate CPU-GPU copy overhead**.
- Containerized inference pipeline with Docker and NVIDIA Container Runtime for **consistent deployment across development and edge production environments**.

Projects

URL Shortener Service | Python, Flask, PostgreSQL, Redis *December 2024 -- Present*

- Built scalable URL shortening service with **custom short-link generation algorithm using base62 encoding**; implemented collision detection and resolution for hash conflicts.
- Designed REST API with PostgreSQL for persistent storage and Redis caching layer to **reduce database queries by 75%** for frequently accessed URLs.
- Added **analytics tracking for click-through rates, geographic distribution, and referrer data**; implemented rate limiting to prevent abuse and expiration mechanism for temporary links.
- Deployed on cloud infrastructure with environment-based configuration management and comprehensive API documentation.

Natural Language to SQL Query System | Python, Flask, PostgreSQL *September 2024 -- November 2024*

- Built REST API service that **translates natural language queries into SQL using LLM integration** and executes them against relational databases; implemented authentication, request validation, and structured error handling for 12+ query patterns.
- Designed safety mechanisms including input sanitization, query validation, parameterized queries to prevent SQL injection, and retry logic for malformed outputs, **improving request success rate by 65%**.
- Containerized application with Docker and documented deployment process for reproducible setup across development and production environments.

Certifications & Publications

- **Oracle Cloud Infrastructure 2025 Generative AI Professional** -- October 2025
- **Oracle Cloud Infrastructure 2025 Data Science Professional** -- September 2025
- **Publication:** Co-authored research on plant disease classification (IGI Global, 2024); **99.61% accuracy**.