Goutham Raju Kosuru Srinivasa

gkosurus@asu.edu | +1 (602) 696-9216 | linkedin.com/in/gouthamraju11 | github.com/Gouthamraju11

Education

Arizona State University, Tempe, AZ — Dec 2025
Master of Science in Information Technology — **GPA: 4.0/4.0**

Technical Skills

Languages: Python, Java, JavaScript/TypeScript, C#
Framework/Backend: React.js, Angular, .NET Core, Node.js, FastAPI, Django, REST APIs, Microservices
Cloud/DevOps: AWS (Lambda, EC2, S3, DynamoDB), Azure (App Services, SQL, DevOps), Docker, CI/CD
Data/ML: TensorFlow, PyTorch, LangChain, SQL/NoSQL (MySQL, MongoDB, DynamoDB), Pandas, NumPy

Experience

Neudesic (An IBM Company) — Associate Software Engineer — Apr 2022 – Nov 2023
Tech: .NET Core, Angular, React.js, Azure, Node.js, Python, TensorFlow, REST APIs, xUnit, Jest

- Designed and deployed a large-scale field-service management system for a **Fortune 500 client** using Angular and .NET Core, serving **2,000+ daily active users**.
- Contributed to **predictive caching implementation** using Python data analysis to identify usage patterns, helping reduce server load and improve response times for frequently accessed resources.
- Architected **microservice modules for high availability** using Azure App Services with sharded Azure SQL databases, enabling distributed transaction processing and reliable concurrent operations.
- Optimized backend caching layers and database query patterns, improving **API throughput by 25%** and reducing **page load latency by 30%** through strategic indexing and Redis integration.
- Built **CI/CD pipelines** with Azure DevOps using YAML-based automation and parallelized job execution, reducing **deployment time by 60%** and increasing **release frequency by 40%**.
- Enhanced code reliability by implementing dependency injection patterns and expanding **unit test coverage from 25% to 75%** using xUnit and Jest, reducing post-deployment regression issues.
- Implemented **distributed tracing** with Azure Application Insights and Log Analytics for real-time performance monitoring, cutting **production issue resolution time by 30%**.

The Sparks Foundation — Software Engineer Intern — Aug 2021 – Oct 2021
Tech: AWS CDK, Lambda, API Gateway, DynamoDB, Python, Docker, CloudWatch

- Developed a **serverless transaction monitoring API** using AWS Lambda, API Gateway, and DynamoDB with event-driven architecture, handling high-volume concurrent requests with consistent low-latency performance.
- Automated **infrastructure provisioning with AWS CDK** (Infrastructure-as-Code), integrated with CodePipeline and CodeBuild, reducing **environment setup time from 2 hours to under 20 minutes**.
- Containerized services with **Docker** for reproducible builds and CI integration, improving deployment consistency across development, staging, and production environments.
- Configured CloudWatch alarms and structured logging for **real-time monitoring and automated alerting**, improving fault tolerance during concurrent Lambda invocations.

Projects

Multi-Agent AI Workflow Automation Platform — Aug 2025
Tech: AWS Bedrock, SageMaker, LangChain, CrewAI, FastAPI, React.js, FAISS

- Designed an intelligent workflow system using **LangChain and CrewAI** for **autonomous data extraction, summarization, and decision-making** across enterprise datasets.
- Integrated **AWS Bedrock and SageMaker** for distributed inference, building a FastAPI backend with **FAISS vector databases** for semantic search.

Book Recommendation System — Dec 2024
Tech: Python, Pandas, NumPy, SciPy, Collaborative Filtering

- Built a **collaborative filtering engine** using cosine similarity on sparse user-item matrices (**60K+ records**) with SciPy operations for memory-efficient computation.
- Designed **ETL pipelines** with pandas optimizing similarity computation using NumPy vectorization and caching for **real-time recommendations**.

Reader and Object Detector for the Visually Impaired — Jul 2022
Tech: Python, TensorFlow, YOLO, OpenCV, Tesseract OCR, Raspberry Pi, Text-to-Speech

- **Co-authored published research paper** on assistive AI using **real-time object detection** and text-to-speech.
- Developed end-to-end computer vision system using **YOLO and Tesseract OCR** on Raspberry Pi.
- Optimized inference pipeline achieving **15+ FPS on embedded systems** through GPU acceleration.

Certifications & Publications

"Reader and Object Detector for Blind" - *International Research Journal of Modernization in Engineering Technology and Science (IRJMETS)*, Volume 04, Issue 07, July 2022 — **Impact Factor: 6.752**
Co-author: Research on real-time assistive AI systems using computer vision and deep learning