Yogith Ramanan

773-550-1620 | yogithramana@gmail.com | https://www.linkedin.com/in/yogith-ramanan-2a8293208/ | https://github.com/YogithR

Education

DePaul University — Chicago, IL, USA
Master of Science in Computer Science — *January 2025 – Present*

Sri Krishna College of Engineering and Technology — Coimbatore, Tamil Nadu, India
Bachelor of Engineering in Computer Science — *July 2020 – May 2024*

Experience

GURU Information Technology Services — January 2024– June 2024
Product Development and Support Engineer Intern — *Chennai, Tamil Nadu, India*

- Debugged customer-facing UI issues across HTML, CSS, and JavaScript by reproducing defects, isolating root causes, and validating fixes end-to-end.
- Investigated production incidents through log analysis, delivered patches, and documented remediation steps to minimize repeat issues.
- Authored detailed defect reports and validation summaries to support engineering resolution and improve cross-functional communication on production issues.

Gateway Software Solution — July 2022 – Sep 2022
Full Stack Intern — *Coimbatore, Tamil Nadu, India*

- Built and enhanced web components that converted backend data into interactive dashboards and user-facing interfaces.
- Maintained front-end, back-end, and database layers to support evolving requirements and reduce regressions.
- Applied GitHub-based workflows for version control and team collaboration to improve review efficiency and deployment readiness.

Paper Publications

Carbon-Aware Scheduling for Multi-Region AI Inference: January 2026 - March 2026
Research Project — Python · Cloud Systems · Simulation · Sustainability

- Designed and implemented a Python-based discrete-event simulation framework evaluating 8 cloud request routing policies across 5 global AWS regions and 3 AI workloads (BERT-base, BERT-large, ResNet-50), processing 33,600 requests over a 7-day window using real-world carbon intensity and latency data.
- Engineered a Constrained Hybrid scheduling policy that achieved 54.8 points of carbon intensity reduction with zero SLO violations, outperforming all static routing baselines while maintaining an average latency of 62.3 ms.
- Built a novel Adaptive Hybrid controller using a closed-loop feedback mechanism that dynamically adjusts routing weight based on real-time P95 latency headroom per workload, eliminating the need for manual parameter tuning.
- Conducted full Pareto trade-off analysis across all policies, identifying optimal deployment conditions for enterprise cloud architects including SLO strictness thresholds, user geographic distribution, and regional carbon contrast requirements.
- Produced an enterprise decision framework mapping 5 deployment scenarios (strict SLO, moderate SLO, relaxed SLO, ultra-strict, and batch async) to recommended policies, with benchmarked comparisons against CASPER, Google CICS, and Microsoft carbon-aware computing systems.

Projects

A2A Java SDK - Agent-to-Agent Communication Platform — January 2025 – June 2025

- Developed a Java SDK client for agent-to-agent communication with typed request/response models and metadata support.
- Implemented streaming response handling and structured logging to improve observability and end-to-end traceability.
- Improved integration reliability by validating payloads, addressing edge cases, and standardizing error handling.

Food Calorie Tracker August 2025 – October 2025

- Developed an image classification pipeline using MobileNetV2 transfer learning on Food-101 with input filtering to reduce non-food requests.
- Deployed a Streamlit-based application and optimized CPU inference through caching and efficient preprocessing to improve latency and runtime stability.
- Integrated human-in-the-loop corrections and logging workflows to improve dataset quality and generate downloadable prediction history.

Serverless Document Processing System (AWS) December 2025 – January 2026

- Architected an event-driven, fault-tolerant document pipeline using S3, Lambda, and asynchronous Textract with DynamoDB-based state tracking for idempotent processing.
- Developed RESTful retrieval endpoints with API Gateway and API-key authentication while monitoring latency and failures through CloudWatch logs and metrics.
- Enforced structured lifecycle handling with 4xx/5xx responses, safe input validation, and least-privilege IAM controls across services.

Cloud-Native Multimodal AI Platform for Drug Candidate Prioritization February 2026 – April 2026

- Built an end-to-end pipeline for drug-like candidates using SMILES, RDKit descriptors, Morgan fingerprints, and metadata, with reproducible ingestion, feature generation, and YAML-driven training workflows.
- Trained and compared logistic regression, Random Forest multimodal fusion, and a PyTorch graph–tabular fusion model (GCN-style graph encoder + tabular MLP); logged experiments in MLflow and produced a structured three-way evaluation across performance, rank agreement, and error analysis.
- Implemented deterministic candidate prioritization from model scores with rule-based descriptor penalties, confidence scoring, tie-breaking, and reason codes to support transparent decision support.
- Deployed a containerized FastAPI service exposing predict, rank, and batch-rank endpoints; automated testing with pytest and GitHub Actions CI; packaged the stack with Docker Compose; and instrumented the API with Prometheus-compatible /metrics and structured request logging.

Technical Skills

Programming Languages: Java, Python, SQL, JavaScript.
Frontend/ UI: React, HTML5, CSS, Streamlit.
Backend/ Integration: REST APIs, FastAPI, OpenAPI, Uvicorn, Flask, Django, Spring Boot, Node.js, Data Integration, Real-Time Data Processing, JUnit, Maven.
Cloud/ DevOps: AWS (EC2, S3, RDS, Lambda, API Gateway, CloudWatch, AWS Batch, VPC, SageMaker, EKS), Docker, Docker Compose, Kubernetes, Terraform, Git, GitHub Actions, CI/CD, Jenkins, Linux, MLflow, pytest.
Platforms and Tools: Databricks, VS Code, PyCharm, Eclipse, IntelliJ IDEA, Jupyter Notebook, MySQL Workbench, Google Colab, YAML.
Databases: PostgreSQL, MySQL, Oracle, MongoDB, SQL Server, DynamoDB.
Software Development Principles: Data Structures and Algorithms, OOP, SDLC, Agile/Scrum, SOLID Principles, Design Patterns, Unit Testing, Integration Testing, Debugging, Production Support, Incident Investigation, API Security, Cloud Security Fundamentals, Reproducible ML, Ranking / Decision Support.
AI/ ML: Machine Learning, Deep Learning, Multimodal AI, Graph Neural Networks (GCN), TensorFlow, PyTorch, scikit-learn, NumPy, Pandas, Feature Engineering, Model Evaluation, Experiment Tracking, Data Preprocessing.
Biotech/ Cheminformatics: RDKit, SMILES, Molecular Descriptors, Morgan Fingerprints, Drug Candidate Prioritization, Molecular Property Prediction.
Visualization: Tableau, Power BI, Plotly, Matplotlib, Excel.