Ramya Ganesh

+1 412-450-2378 ramyaganesh.cmu@gmail.com github.com/Rganeshk linkedin.com/in/ramyapriyanandhini

EDUCATION

Carnegie Mellon University (CMU) | School of Computer Science *Pittsburgh, PA*
Master of Software Engineering *Dec. 2025*
Relevant Courses: Distributed Systems, Cloud Computing, Software Architecture, Machine Learning, Generative AI
2025 MSE Director's Scholarship - *Awarded for leadership, program citizenship, and academic excellence*

Amrita Vishwa Vidyapeetham University *Coimbatore, India*
Master of Technology (Artificial Intelligence & Data Science) *Jul. 2021*

WORK EXPERIENCE

eParts Services (CMU Capstone Project) *Jan. 2025 – Dec 2025*
Senior Software Engineer - Cloud-Native Data Platform | ML Ops
- Led a cross-functional team of 6 in architecting a **multi-tenant cloud-native Data Warehouse** on Azure, harnessing Snowflake for real-time analytics and cutting down reporting latency by **60%** across **∼50K+** records/day across **∼15+** tenants.
- Architected **fault-tolerant Real-time ELT pipelines** (Kafka, Debezium) on Azure Kubernetes Service (AKS) to stream **∼150K+** SQL Server CDC events/day into Snowflake, achieving **85%** improvement in data freshness and enabling scalable ML feature engineering.
- Unified advanced observability using Snowflake query logs, Grafana dashboards, and Prometheus metrics, decreasing incident triage time by **50%** and ensuring **>99.9% SLA** adherence for **∼75+** critical data workflows.

Philips

Software Engineer II - Cloud-Native Medical Imaging Platform | SaaS *Mar. 2023 – Jun. 2024*
- Pioneered a scalable **Medical Imaging Platform** (Annotation Viewer) on AWS EKS, replacing legacy systems with **Kubernetes-orchestrated containers**, facilitating AI-assisted annotation (e.g., auto-segmentation, lesion detection) and boosting radiology efficiency by **∼30%** for **∼15k+** daily sessions.
- Designed **fault-tolerant imaging services** with Angular/TypeScript and .NET Core, using REST and web sockets for low-latency communication across 5+ DICOM modalities (CT, MR, PET, X-ray, US), handling **∼20K+** image requests/day with **zero downtime**.
- Optimized backend performance with AWS ElastiCache and DynamoDB, implementing data pipelines for real-time ML inference, attaining **99.999%** uptime across **∼10K+** diagnostic workflows in multi-region setups.
- Collaborated globally (Netherlands, Israel) to architect HIPAA-compliant data flows, incorporating Kafka for streaming **∼500K+** image metadata events/day, supporting interoperable AI pipelines.
- Mentored 5+ engineers on containerized workflows, streamlining CI/CD with GitHub Actions, cutting onboarding time by **40%** in cross-region teams.

Software Engineer - Platform Engineering | SaaS *Aug. 2021 – Mar. 2023*
- Revamped legacy imaging into a modular SaaS platform with C#/.NET Core and Angular, leveraging MongoDB sharding to reduce latency by 30% for **∼8K+** daily diagnostic sessions.
- Accelerated deployments using Docker and AWS ECS, automating Terraform-based CI/CD pipelines, accomplishing **99.99%** release reliability and slashing cycle times by **80%** across regions.
- Implemented caching and streaming with AWS Kinesis, cutting image load times by **40%** and powering seamless access for **∼5K+** concurrent users in high-throughput platforms.

Software Engineer Intern – Data Engineering Team *Aug. 2020 – Jul. 2021*
- Built scalable pipelines with Apache Spark and Kafka, processing **∼1M+** healthcare events/day to power ML-driven anomaly detection, enhancing clinical system insights.
- Developed a health check API with ReactJS, Java, and Spring Boot, using Consul for service discovery, boosting node uptime from **95% to 99.99%** across **∼10+** nodes.
- Automated **65%** of testing with Python scripts in CI/CD pipelines, ensuring robust data processing for healthcare analytics workloads.

PROJECTS

NYCabs - Ride-Matching and AI-Powered Travel Agent | Cloud Computing CMU *Sep. 2025 – Sep. 2025*
- Engineered Kafka-Samza pipeline to stream driver-rider events with real-time driver assignments, achieving ∼**96%** match accuracy.
- Created XGBoost fare prediction model with engineered features (distance, anomalies), tuned on Vertex AI achieving 3.58 RMSE.
- Deployed RAG-powered travel agent using Vertex AI RAG Engine and Llama-3.1-70B-Instruct for **90%** relevant NYC query response.

Multi-Object Cardinality Preservation in Stable Diffusion | Generative AI CMU *May. 2025 – May. 2025*
- Extended **Stable Diffusion** for multi-object count fidelity in text-to-image generation by adopting Slot-Guided Attention and ALiBi embeddings to better enforce spatial layout and object relationships.
- **Fine-tuned** with LoRA on structured cardinality datasets, securing a **31%** F1-score improvement and **0.93** precision on multi-object generation tasks with spatial layout constraints.

SKILLS

Languages and Frameworks: Python, C#, .Net, Angular, React, Typescript, Node.js, Java, Playwright, Bash, Powershell, FastAPI.
Databases and Data Processing: MongoDB, Elasticsearch, Kafka, Samza, Hadoop, Hive, Pig, Spark, Redis, MySQL.
Cloud and DevOps: AWS (EC2, S3, IAM, VPC, Lambda), Azure, GCP, Helm, Jenkins, Git, Terraform, Docker, Kubernetes.
Machine Learning and AI: Pytorch, pandas, XGBoost, LangChain, Vertex AI, Cursor, Huggingface, Transformers, LLM, RAG, MCP.
Leadership and Project Management: Agile (Scrum, Jira), Quality Assurance, Requirement Analysis, TDD, Product Management.