

Lohit Geddam

925-364-1402 | lohit.geddam@gmail.com | linkedin.com/lohitg | github.com/lohitg

EXPERIENCE

Technical Co-Founder

Sept 2025 – Present

Tandem - Backed by Foundation Capital

San Francisco, CA

- Built “Gardener,” an autonomous Slack-to-PR coding agent powered by Claude that converts engineering requests into production-ready pull requests with automated code review and iteration workflows
- Built a production tool-calling AI agent for engineering workflows with semantic search, CRUD operations, and persistent multi-turn context, reducing manual workflow steps by 80%
- Built a requirements management system with hierarchical traceability, baseline versioning, custom attributes, saved views, and CSV/Excel import, integrated with pgvector semantic search across 20+ PostgreSQL tables
- Authored the full AWS infrastructure in Terraform across commercial and GovCloud environments, including VPC, ECS Fargate, SQS workers, RDS pgvector, Cognito, CloudFront, and WebSocket Lambda APIs
- Architected a multi-tenant FastAPI + React/TypeScript platform with Cognito authentication, org-scoped isolation, and WebSocket-based live collaboration

Software Development Engineer

Sept 2024 – Sept 2025

Amazon Web Services (AWS) - AWS SageMaker HyperPod

Bellevue, WA

- Engineered scalable ML infrastructure capable of managing 10k+ GPU superclusters orchestrated by Kubernetes
- Developed high-concurrency features in Go for key workflows, enabling efficient distributed ML workloads at scale
- Designed a feature for pre-caching ML training images across large-scale clusters to reduce MTTR by 33%
- Resolved cluster issues for customers using kubectl troubleshooting and Prometheus/Grafana observability tools
- Created a full CI/CD pipeline in TypeScript using Infrastructure as Code (IaC), incorporating integration tests and canary deployments with automated rollbacks to continuously deliver features to customers
- Optimized node movement between clusters, reducing transfer time by 83%, significantly improving resource utilization between large-scale ML workloads

Software Development Engineer Intern

June 2023 – Sept 2023

Amazon - Alexa Kitchen

Sunnyvale, CA

- Designed and developed a new feature on the Alexa Mobile App using Typescript and React Native allowing over 100 million customers to discover new recipes through personalized notifications using LLMs with A/B Testing
- Worked with multiple engineering teams and stakeholders to identify trade-offs and key feature requirements
- Implemented a new back-end server using Flask and Python to reduce recipe retrieval latency by 80%

Software Development Engineer Intern

June 2022 – Sept 2022

Amazon Web Services (AWS) - AWS Data Platform

Seattle, WA

- Developed and automated PySpark jobs to clean and profile large amounts of data (600GB+) through AWS Glue
- Updated project documentation per design and participated in scrum/agile meetings to update project progress
- Developed a new dashboard using React.js to allow customers to easily access data metrics for 1,000+ tables

PROJECTS

Fake News Detection | *Google Cloud, Vertex AI, Python, LLMs, Generative AI, NLP*

Sept 2023 – Mar 2024

- Developed predictive multi-class models with sklearn using 25,000+ articles to rank news article truthfulness
- Employed Google Vertex AI and AutoML to deploy additional predictive models for real-time predictions
- Utilized Google Gemini AI with RAG and Vector Databases for feature engineering and fact checking
- Combined predictive multi-class models with Generative AI outputs to create a cohesive article ranking system

TECHNICAL SKILLS

Languages: Python, Go (Golang), TypeScript, Java, JavaScript, SQL, R, HTML/CSS, Shell Script

Frameworks/Tools: React, Kubernetes, Docker, Terraform, PostgreSQL, Spark, FastAPI, Git, AWS, pgvector, REST

Skills: Machine Learning, AI Agents, CI/CD, ML Infrastructure, Backend Systems, Mobile & WebApp Development

Certifications: [AWS Certified AI Practitioner](#)

EDUCATION

University of California, San Diego

B.S. in Data Science, Minor in Marketing

La Jolla, CA

June 2024