Gokul Asamani

linkedin.com/in/gokulprathin
Fairfax, Virginia

Email : gokulprathin8@gmail.com
Mobile : +1-571-297-5966

Summary

Highly motivated Full Stack Software Engineer and Machine Learning practitioner with over 4 years of experience designing, developing, and deploying scalable cloud-based applications and AI systems. Proven expertise in full-stack development (Python, React/NextJS, Node.js), robust system architecture, and cloud platforms (AWS, GCP). Adept at building reliable, high-performance solutions using agile methodologies. Extensive experience with Generative AI, including Large Language Models (LLMs like Llama 7B/70B), Retrieval-Augmented Generation (RAG) systems using LangChain, vector databases (FAISS, ChromaDB), model fine-tuning (LoRA/QLoRA, Quantization), model deployment (vLLM), and developing AI-driven applications with multimodal capabilities using frameworks like CrewAI and models like Google Gemini/Gemma. Consistently focused on enhancing performance, user engagement, and leveraging AI for innovative solutions.

Experience

Full Stack Software Engineer *09/2024 – 04/2025*
Agrograph *Fairfax, USA*

- Architected and implemented scalable cloud infrastructure on Google Cloud Platform (GCP) using container orchestration with Google Kubernetes Engine (GKE) to support Agrograph's geospatial data processing platform, demonstrating strong system design practices.
- Established and managed CI/CD pipelines using GitHub Actions, automating build processes, testing, and deployment to ensure platform reliability and adherence to coding standards.
- Collaborated across the full software development lifecycle, developing customer-facing features with React/NextJS on the front end and Python on the back end, ensuring system reliability and proper application scaling.
- Implemented and optimized a Tegola vector tile server integrated with PostGIS to deliver performant map visualizations for parcel, tillage, and cover crop data, applying effective design patterns.

Machine Learning Research Assistant *02/2023 – 08/2024*
George Mason University *Fairfax, USA*

- Contributed to the Geoweaver Project, an open-source workflow management tool funded by NASA, enhancing its functionality and improving user engagement through collaborative development efforts.
- Integrated Llama 7B model into Geoweaver to enable local, LLM-assisted code generation for enhanced user workflow creation.
- Deployed a Llama 70B model using optimized inference serving techniques to provide multi-user LLM capabilities within the Geoweaver platform for production.
- Maintained the existing codebase and implemented new features in line with best practices in the software development lifecycle and system design principles.
- Assisted a professor in writing research papers on topics including 'Pygeoweaver Tangible Workflows' and 'A Review of Practical AI for Remote Sensing in Earth Sciences', reinforcing technical documentation skills.
- Created an open-source Python bindings library for Geoweaver (github.com/ESIPFed/pygeoweaver) to facilitate advanced analyses and improve accessibility for the research community.
- Developed a Snow Water Equivalent workflow utilizing SRTM DEM and Gridmet data to predict water in snow across the Western US, applying robust software engineering and system reliability techniques.

Software Engineer *06/2021 – 08/2022*
Marvin (fka Userfocus) *Telangana, India*

- Developed full-stack applications using Django, Flask, AWS, Elastic Search, Postgres, Heroku, React, Redux, and Webpack, enhancing system performance and user experience.
- Closely worked with product team to create dashboards on Retool for analysis and gathering insights on how users interact with the platform for laying down future roadmap for the company.
- Worked on Zoom integration into the existing codebase and built an end-to-end pipeline to record and analyze calls from MS Teams & Google Meet, automating the recording process with Selenium, FFmpeg & Chrome headless.

Associate Software Engineer *12/2019 – 05/2021*
Offerly *Hyderabad, India*

- Worked on building a full stack E-Commerce application. Built a functional administration panel for active monitoring of customers and order details.
- Implemented business requirement features in production using React, Redux, Postgres, AWS, Node.js, and Sentry, improving application functionality and reliability.
- Responsible for migrating the project from React to NextJS for better performance and server-side rendering.

Education

George Mason University *Fairfax, Virginia*
Master's degree in Computer and Information Sciences

GITAM Deemed University
Bachelor of Technology in Computer Science

Projects

Multimodal Generative AI Applications: Developed a suite of end-to-end applications leveraging Google Gemini (Pro/Vision) and Gemma models. Implemented conversational AI chatbots, document/PDF Q&A systems using LangChain and RAG techniques, multimodal analysis tools (image/voice processing), and a Text-to-SQL application for database querying. Explored model fine-tuning (LORA/QLORA) with custom datasets for specialized tasks like ATS resume analysis and summarization.

Autonomous AI Agent Workflows: Built multi-agent systems using CrewAI framework integrated with LLMs (Gemini Pro, OpenAI API). Designed and implemented agents capable of autonomous task execution, such as a news reporter agent for information gathering/summarization and a content creation agent transforming YouTube video content into blog posts. Focused on collaborative agent workflows and automation.

Geoweaver (02/2023 – 04/2025): In-browser software for composing and executing full-stack data processing workflows, leveraging online spatial data, HPC platforms, and open-source deep learning libraries. Provides server management, code repository, workflow orchestration, and history recording.

Pygeoweaver (02/2023 – 04/2025): Python wrapper for the Java-based GeoWeaver app, enabling direct use within Python, Jupyter Notebook, JupyterLab, and JupyterHub environments for enhanced research accessibility.

Snowcast (05/2023 – 03/2024): Developed a Snow Water Equivalent (SWE) prediction workflow using Geoweaver. Utilized Python and Shell scripts to collect GridMET meteorology data. Trained an AI-based model (potentially leveraging ML/DL techniques) for daily, 4-km spatial resolution SWE prediction across the western US.

Natours: Tour booking application built with the MERN stack (MongoDB, Express.js, React/Redux, Node.js). Features user authentication with JWT, role-based access control, and secure password management.

Skills

AI/ML: Generative AI, LLMs (Llama 7B/70B), RAG, LangChain, Vector DBs (FAISS, ChromaDB), Fine-tuning (LoRA, QLoRA), Quantization, Model Deployment (vLLM, TGI), Transformers, Neural Networks, Deep Learning, ML, Hugging Face, Google Gemini (Pro, Vision), Google Gemma, CrewAI, Multimodal AI, AI Agents, MLOps, PyTorch, Python (ML Libraries), Geospatial AI, API Integration (OpenAI, Groq)

Programming: Python (Django, Flask), JavaScript/TypeScript (React, NextJS, Node.js, Express.js), Java, SQL, HTML5, CSS3 (Tailwind CSS), Redux, Webpack

Cloud & DevOps: GCP, AWS, GKE, Docker, CI/CD (GitHub Actions), Heroku, Serverless, Microservices

Databases: PostgreSQL (PostGIS), MongoDB, Elasticsearch, SQL, NoSQL

Tools & Methodologies: Git, Agile, System Design/Architecture, REST APIs, GraphQL, Selenium, FFmpeg, Retool, Sentry, JIRA, Technical Documentation, Code Reviews, Testing, GIS

Publications

MDPI. A Review of Practical AI for Remote Sensing in Earth Sciences: Synthesized and analyzed AI methodologies and outcomes in remote sensing, identifying research gaps and emerging trends. Explored diverse applications including image classification, land cover mapping, and data fusion, while addressing challenges like data quality and model interpretability.

Certifications

Oracle Cloud Infrastructure 2024 Generative AI Certified Professional

Microsoft Power BI Desktop for Business Intelligence – Maven Analytics

Python Data Structures – Coursera

Python Geospatial Navigation System – ISRO

Geographic Information Services Professional (GIS) – ISRO

PyTorch for Deep Learning – Pierian Training