Sangeet Saha

Boston, MA, 02120

+1-857-693-1330 | sangeetsaha1116@gmail.com | https://linkedin.com/in/sangeet-saha-a2b1651aa

Summary

Full-stack developer and cloud enthusiast with experience building scalable web applications and AI-powered systems. Proficient in the MERN stack, CI/CD, and GCP infrastructure. Delivered platforms integrating LLMs, RAG pipelines, and vector databases. Seeking full-time roles in software development or cloud engineering to build innovative, reliable systems at scale.

Technical Skills

- **Languages:** C++, Java, Python, Javascript, Bash
- **Web Technologies:** HTML, CSS, React, Node.js, Spring Boot
- **Cloud & DevOps:** Git, GitHub Actions, Terraform, Packer, GCP
- **Databases:** MySQL, MongoDB, PostgreSQL, Pinecone, ChromaDB, Qdrant.
- **Other Tools:** Linux, ANSYS, SolidWorks, Figma, Postman, Trello, Moquups, Ollama, Neo4j, Docker
- **Software Engineering:** Distributed Systems, Object-oriented Design
- **Interpersonal Skills:** Excellent Communication, Presentation skills, Analytical skills, Collaboration

Professional Experience

Northeastern University | *Teaching Assistant: Agile Software Development*   **May 2025 - Present**

- Facilitated agile simulations and group activities for 72 students, enhancing iterative learning and collaboration.
- Managed a Slack channel to resolve student queries, promoting effective communication in a dynamic learning environment.
- Developed and graded assessments using established question banks, ensuring fair and consistent evaluation.
- Administered the Canvas course page with weekly modules, assignments, and announcements to ensure smooth course delivery.
- Collaborated in weekly meetings with the professor to review progress, refine content, and plan upcoming instructional materials.

Humanitarian AI | *AI Engineer*   **Jan 2025 - April 2025**

- Engineered a chain-of-thought AI architecture using Python to address challenges in shallow reasoning.
- Architected tree-structured AI agents to decompose complex tasks into specialized subtasks, enhancing workflow efficiency.
- Integrated tools such as Gradio, LangChain, Ollama, and Neo4j to execute end-to-end AI solutions in a scalable environment.
- Collaborated with cross-functional teams and mentored teammates on emerging AI technologies and best practices.
- Streamlined the AI application development lifecycle from design to deployment to support scalable system builds.

Wipro | *SAP Basis Consultant*   **Jul 2021 - Aug 2023**

- Optimized SAP ECC/S4HANA systems, enhancing performance by 35% and monitoring 1,000+ daily transactions.
- Led maintenance, troubleshooting, and client reporting while mentoring colleagues to reduce resolution time by 20%.
- Streamlined intern hiring by coordinating 50+ interviews and improving feedback processes for greater efficiency.

Education

Northeastern University
Master Of Science, Software Engineering Systems (GPA: 3.8/4)
    Coursework: SDLC, Cloud Computing, Agile Software Development, AI Infrastructure

Vellore Institute of Technology
Bachelor Of Technology, Mechanical Engineering
    Coursework: C++ Programming, Data Structures, and Algorithms

Relevant Projects

REST API Development and GCP Infrastructure Setup *(Jan 2024 - Mar 2024)*

- Developed and deployed a Node.js REST API with user authentication, utilizing Sequelize ORM and MySQL.
- Implemented CI/CD pipeline with GitHub Actions and Jest for integration testing.
- Utilized GCP for infrastructure, including Cloud SQL, load balancing, and Cloud DNS.
- Automated environment setup using Terraform (IaC) and Packer for custom images.
- Enhanced security with encryption key rings and implemented logging with Winston to GCP Log Explorer.
- Integrated emailing and notification services for improved user communication.

Fair Trader AI: LLM-Powered Online Marketplace *(Oct 2024 - Nov 2024)*

- Designed and implemented an intelligent marketplace using fine-tuned Llama3.2:3B for dynamic price negotiation.
- Integrated Retrieval-Augmented Generation (RAG) for pricing trends, user preferences, & context-aware interactions.
- Achieved 24/7 uptime, reducing operational costs by 50–80%.