Ayush Gupta

LinkedIn: itsayush-live
Github: github.com/ayushgml
More at: itsayush.dev
Mobile: +91-9368980677
Twitter: itsayush

Skills

- **Programming**: Python, Bash, JavaScript, SQL; PyTorch; TensorFlow; Autogen; CrewAI; LangChain; Google ADK
- **Cloud Technologies and Tools**: Amazon Web Services (AWS); Google Cloud Platform (GCP); Azure Foundry; Docker; Cloud Computing; Infrastructure as Code (IaC); Claude; Claude Code
- **AI/ML Expertise**: AI Agents & Agentic Orchestration; Harness Engineering; Retrieval-Augmented Generation (RAG); Agentic-RAG; Model Context Protocol (MCP); Generative AI; Natural Language Processing; Deep Learning; Prompt Engineering; Responsible AI & AI Ethics
- **AI Strategy and Business**: AI Strategy; GenAI Adoption; Use-Case Identification; Cross-Functional Collaboration

Experience

- **Accenture** — Bangalore
  Full Stack LLM Development Analyst — *June 2024-Present*
  - **Agentic AI**: Developed and deployed a multi-agent AI architecture that automated complex decision flows and reduced manual oversight by 50%.
  - **RAG Systems**: Built enterprise-grade RAG systems that grounded LLMs in proprietary data, cutting hallucinations by 50% and raising contextual accuracy.
  - **MCP & Tool-Use Architecture**: Architected Model Context Protocol (MCP) server integrations enabling AI agents to interact with enterprise APIs and internal tooling, reducing custom glue-code by 40% and accelerating agent capability expansion.
  - **Vector Database & Embeddings**: Implemented scalable vector search pipelines using embedding models and semantic chunking strategies, improving retrieval precision by 35% and reducing end-to-end query latency for RAG workloads.
  - **GCP ML Infrastructure**: Provisioned and managed GCP Vertex AI workspaces and Cloud Run services for model serving, enabling zero-downtime deployments and auto-scaling inference endpoints that handled 3x traffic spikes without manual intervention.
  - **Telecom Client Integration**: Integrated AI agents into a major Indian telecom cloud ecosystem and created Autogen-based cloud wrappers, improving operational workflows by 25%.
- **BBIOT Technologies** — Hybrid (Vellore)
  Software Development Intern — *Feb 2023 - Apr 2023*
  - **Backend Development**: Built 25+ NodeJS/Express APIs (4000+ LOC) serving 10,000+ IoT sensors across real-estate microcontroller devices, supporting production features and internal tooling.
  - **IoT Collaboration**: Optimised Python APIs with async request handling and batching for low-cost legacy edge devices, boosting integration reliability by 25% and cutting communication latency by nearly 50%.

Education

- **Vellore Institute of Technology** — Vellore, India
  B.Tech - CSE (Bioinformatics); CGPA: 8.67 — *Oct 2020 - May 2024*

Courses: *Data Structures, Analysis Of Algorithms, Computer Networking, Operating Systems, Database Management Systems*

Certifications

- **Claude Certified Architect - Professional (Anthropic)**: Issued October 2026
- **GCP Associate Cloud Engineer Certification**: Issued October 2025
- **GCP Generative AI Leader Certification**: Issued August 2025
- **Oracle Cloud Infrastructure 2024 Generative AI Professional**: Issued July 2024
- **AWS Certified Solutions Architect Associate**: Issued April 2024
- **Microsoft Certified Security, Compliance and Identity Fundamentals**: Issued June 2023

Honors and Awards

- Received the Growth Catalyst ACE award at Accenture for outstanding performance on critical, time-sensitive deliverables
- Selected as a speaker for Cloud Native Rejekts Paris 2024 conference - Feb, 2024
- Selected for Amazon Machine Learning Summer School - July, 2022
- Zeeve track winner at BLOCKET 2022 among 300+ participants - July, 2022
- Second Runner-up at WEHACK (hackathon) among 60+ teams - April, 2022
- Runner's Up at ECO-HACK (hackathon) among hundreds of participants - Nov, 2021