

Vaishal Malu K

AI Solutions Engineer | Generative AI · Agentic AI · RAG · LLM Solutions Architecture
Villupuram, Tamil Nadu, India | Open to relocation – London, UK | +91 9345671593 | intellibytevaishal@gmail.com
LinkedIn | GitHub

Professional Summary

AI Solutions Engineer with hands-on experience designing, building, and deploying LLM-powered, Agentic AI, and RAG-based platforms end to end – from prompt engineering and LLM API integration through to production deployment and customer-facing delivery. Skilled at technical discovery, rapid solution prototyping in Python, and translating business requirements into working AI workflows across healthcare, security, and recruitment. Experienced in multi-agent architecture and cloud deployment on AWS and Azure. Published researcher, patent contributor, and 10+ time national hackathon winner known for building and demoing working AI products at speed.

Technical Skills

AI / LLM & Agentic Systems: Generative AI, Agentic AI, RAG Pipelines, Multi-Agent Architecture, Prompt Engineering, LLM API Integration (Groq, OpenAI-compatible APIs), MCP Orchestration, Foundation Models, AI Accuracy Validation

Programming & Backend: Python, FastAPI, Flask, React.js, HTML, CSS

Cloud & DevOps: AWS, Microsoft Azure, Docker, Git, n8n Workflow Automation

Data, Integrations & Analytics: Firebase, Power BI, Tableau, REST API Integration

Solution Design & Delivery: Technical Discovery, Solution Prototyping, POC Delivery, Stakeholder Communication, Figma, Canva

Experience

Cloudcouch Tech Private Limited

Oct 2025 – Apr 2026

Software Engineering Intern

- Partnered with product and business teams on market research and go-to-market strategy for Wynk Wash, translating customer needs into technical feature requirements.
- Designed and implemented a WebRTC-based in-app calling solution to solve a real customer communication workflow, and owned the Wynk Wash Admin Panel end to end for service operations and workflow automation.
- Led digital engagement and platform management initiatives that directly supported customer acquisition and product adoption.

CubeAI Solutions Tech Pvt Ltd

Mar 2025 – Jun 2025

AI/ML Intern

- Built AI-powered recruitment automation combining NLP, computer vision, and sentiment analysis for intelligent candidate evaluation; designed backend AI workflows and automation pipelines in Python and Flask, delivered as scalable, production-grade enterprise applications.

Google for Developers & India Edu Program

Oct 2024 – Dec 2024

AI/ML Virtual Intern

Built and evaluated machine learning models in Python and TensorFlow on real-world datasets, applying feature engineering, model training, and evaluation pipelines.

Selected AI Projects

MediScribe AI – Enterprise Multi-Agent Clinical Documentation Platform

AI Architect · Backend Developer

- Architected a multi-agent, RAG-based healthcare AI system converting doctor-patient conversations into validated clinical documentation (SOAP notes), with full audit trail and transparent reasoning at each stage.
- Integrated LLM APIs (Groq), FastAPI, MCP-based agent orchestration, and observability tooling; built validation guardrails backed by 119+ automated tests to ensure accuracy in a regulated, high-stakes domain.

Agentic AI Resume Filter & Interview System

AI Developer

- Built an end-to-end agentic AI workflow automating NLP-based resume screening, AI-conducted interviews, and candidate evaluation with emotional-analysis models.

NeuroVibes – AI Assistive Communication Platform

AI Developer · Hardware Integration

- AI-powered assistive communication system with vibration hardware, speech enhancement, and deep learning; validated with 125+ users across 5 special schools, 35% clarity improvement (Patent App. No. 202541055359).

Hackathons & Recognition

- Winner of 10+ national/institutional hackathons in AI, ML, and Cybersecurity (incl. Pitch-A-Thon 2024, VIT Chennai; Innovators Day 2024) | 3rd Place, National Cyber Challenge (MHA/BPRD/NCRB) | Young Innovator Award, Zoho.

Publications

- AI-Driven Energy Forecasting with Generative Analytics & Anomaly Detection for Smart Grid Management* – IJMRSET, Vol. 9, Issue 4, Apr 2026.

Certifications

AWS Academy Graduate – Machine Learning Foundations | AWS Academy Graduate – Cloud Foundations | Oracle Advanced PL/SQL Developer Certified Professional | Data Analytics Essentials, Cisco Networking Academy | Signal Processing, ML, Deep Learning & Image Processing – MathWorks

Education

Sri Manakula Vinayagar Engineering College

2024 – 2027

B.Tech, Artificial Intelligence and Data Science – CGPA 7.49/10 (Ongoing)

Elumalai Polytechnic College – Diploma in Computer Engineering, First Class 76%

2021 – 2024