Rahul Yadav
[GitHub](#) | [Portfolio](#) | [LinkedIn](#) | [Email](#)

PROFILE SUMMARY

Full-Stack & Applied AI Engineer with a deep foundation in computer science and hands-on experience architecting scalable, cloud-native systems. Built and deployed end-to-end products integrating **LLMs, vector search**, and **predictive analytics** across healthcare and SaaS domains. Proficient in **Next.js, Python, AWS, and CI/CD pipelines**, delivering sub-500 ms APIs, 99.9 % uptime, and production-grade performance for real users.

SKILLS

- **Languages:** Languages: Python, JavaScript, Java, TypeScript, C++ (academic), Bash
- **Frameworks:** Next.js, FastAPI, Node.js, React, Tailwind CSS
- **Cloud & DevOps:** AWS (Lambda, EC2, S3), GCP (Cloud Run, App Engine), Docker, NGINX, GitHub Actions, Terraform
- **Databases:** PostgreSQL, DynamoDB, Supabase, Vector DBs (Pinecone, FAISS)
- **Concepts:** System Design, Microservices, CI/CD, Distributed Systems, Data Structures & Algorithms

---

Engineering Experience

Software Engineering Intern — GoRoots Technologies   *Oct 2025 – Present*
- Building and optimizing **backend APIs** for a salon management platform used by customers and businesses.
- Reduced API latency by **40 %** through query optimization and caching strategies.
- Implemented **CI/CD pipelines** and automated error tracking to accelerate deployment cycles.
- Designed **IP-based location suggestions** and developed **vector databases–driven recommendation agents**.

Software Engineer (Volunteer) — Rural HealthTech Project, Nepal   *Sep 2023 – Jan 2025*
- Developed a **React Native + Sheets API** mobile platform to digitize rural patient data, increasing efficiency by **60 %**.
- Built a **serverless backend** using Google Drive API + Sheets API to securely store patient data at zero cost.
- Coordinated logistics and technical strategy to scale operations for **1,000+ local residents**.

Projects

AeroStream – AI-Integrated Health Analytics Platform (Patent Applied)   [**[Link](#)** | **[Code](#)**]
- Engineered an **end-to-end health analytics platform** integrating **Next.js web app**, **Gemini AI**, and custom sensor hardware.
- Designed complete doctor,patient,admin management, real-time analytics, and predictive health monitoring workflows.
- Built custom **AI models for patient-state prediction** and integrated Gemini APIs for clinical report summarization.
- Deployed on AWS with **Docker + CI/CD**, achieving < 500 ms API latency and **99.9 % uptime**.

Uni-Papers — LLM-Powered Academic Platform   [**[Link](#)** | **[Code](#)**]
- Built a production-grade **OAuth 2.0** authentication system supporting **300 + users** with secure login and session flows.
- Integrated **Gemini LLM API** for context-aware Q&A, increasing user session time by **40 %**.
- Architected distributed microservices with load balancing and caching layers, deployed on Vercel with monitoring and < 500 ms API response.

Early Disease Detection via Nail Image – AI-Powered Web System   [**[Link](#)** | **[Code](#)** | **[Demo](#)**]
- Designed and deployed an end-to-end Next.js + FastAPI diagnostic web app leveraging **CNNs for disease prediction.**
- Achieved **<2 s inference latency and optimized ML backend deployment on Render, mitigating cold-start delays in free-tier environments.**
- Engineered REST APIs for real-time predictions and structured logging with alerting pipeline.
- Integrated scalable microservice architecture for inference and data management.

Technical Highlights

- Deployed 6+ production-grade systems across AWS, GCP, and Vercel with 99.9 % uptime.
- Integrated LLMs and vector databases (Gemini, Pinecone) into real-world apps for personalized AI analytics.
- Reduced backend response latency from 1.2 s → <500 ms via optimized architecture and caching.
- Delivered 70% cost reduction in compute via serverless optimization (AWS Lambda + Docker).

Education

B.Tech in Computer Science and Engineering
Vellore Institute of Technology (VIT Vellore)   *[July 2026]*
CGPA 7.96 / 10
 Relevant Coursework: Data Structures & Algorithms, Operating Systems, Database Systems, Computer Networks

---

CERTIFICATIONS / Badges

- **AI Certification**, Smartbridge x Google   [**[Link](#)**]
- **Google** Cloud Skills Boost Badges (3+ badges)   [**[Google Badges](#)**]
- **IBM** Skill Build Badges (2+ badges)   [**[IBM Badges](#)**]