

Rithwik Marpally

rithwikmarpally07@gmail.com — +91-8143962706 —  github.com/rithu453 —  linkedin.com/in/rithwik-marpally — rithwik.social

Technical Skills

Languages: C++, Python, Java, JavaScript
Frameworks: FastAPI, Node.js, MERN Stack
Cloud/Infra: AWS, Docker, Kubernetes, CI/CD
AI/ML: LLMs, RAG, Transformers, TensorFlow
Systems: Distributed Systems, Microservices, Data Structures & Algorithms

Experience

Feuji — Hyderabad, India May 2025 – Aug 2025
AI Engineer Intern

- Engineered AI agents using Model Context Protocol (MCP), Python, and Playwright to automate cross-environment test execution, cutting QA time by 70%.
- Optimized FastAPI and Node.js microservices, improving average API latency by 35% and throughput by 20%.
- Designed REST APIs integrated with Electron + React apps, eliminating 40% of manual QA workflows.
- Supported deployment of microservices on AWS by configuring Docker containers and CI/CD workflows under senior guidance.

Projects

Samvaad — AI Interview Assistance Platform — MERN, FastAPI, VLM, AWS GitHub

- Built an adaptive AI interview platform with automated question generation and code evaluation.
- Integrated Vision-Language Models (VLMs) for behavioral and gesture analysis with 85% accuracy.
- Deployed containerized services and async task queues on AWS to achieve sub-second response latency.
- Tested with 100+ real users across college; increased interview evaluation throughput by 60%.

Samarth AI — AI-Powered Test Automation Platform — FastAPI, Playwright, MCP, AWS GitHub

- Automated 50+ cross-environment web test cases using MCP, reducing QA effort by 60% and execution time by 70%.
- Architected distributed caching and asynchronous pipelines for fault-tolerant scalability.
- Deployed end-to-end system on AWS with containerized microservices, CI/CD, and centralized monitoring.

LAW Assistant — Legal Information Platform — MERN, FastAPI, RAG, Vercel GitHub

- Engineered a scalable legal information platform simplifying access to common laws and legal resources for the public.
- Designed and deployed a RAG-based chatbot using BERT embeddings, delivering 92% accurate responses with around 10s latency.
- Implemented semantic retrieval for legal research materials and case references to enhance query precision.
- Deployed on Vercel for high availability and sub-second load times; adopted by many users for legal learning and research.

Education

Keshav Memorial Institute of Technology — B.Tech in Computer Science 2022 – 2026
CGPA: 8.0

Relevant Coursework: Data Structures & Algorithms, Operating Systems, Database Systems, Cloud Computing, Machine Learning, Distributed Systems

Achievements & Contributions

- 1st place — IIIT Sri City Hackathon (100+ teams) for AI-driven automation platform.
- Top 10 finalist — IIT Bombay Pulse Hackathon (AI domain) among 200+ teams.
- Top 5 — Feuji Hackathon (20+ teams) for scalable AI integration.
- Developed an interpreted programming language in Sanskrit named **Paanini**, exploring symbolic reasoning and computational linguistics.
- Volunteer at National Service Scheme (NSS) and member of Rotaract KMIT.