

ROHAN PRAVEEN CHAVAN

Blacksburg, VA | +1-540-824-9961 | rohanchavan0701@gmail.com
linkedin.com/in/rohanpraveenchavan | github.com/RohanChavan0701 | rohanchavan.vercel.app

EDUCATION

Virginia Polytechnic Institute and State University

M.S. Computer Engineering, GPA: 3.71/4.0

Aug 2024 – May 2026

Blacksburg, VA

– **Coursework:** Advanced Machine Learning, Computer Vision, Trustworthy ML, IoT System Design

K.J. Somaiya College of Engineering

B.Tech Information Technology (Distinction; Honors in AI)

Aug 2020 – May 2024

Mumbai, India

TECHNICAL SKILLS

Languages: Python, TypeScript, SQL, C++

Backend: FastAPI, Node.js, REST APIs, microservices

AI/ML: PyTorch, OpenAI/Claude APIs, RAG pipelines, prompt engineering

Infrastructure: AWS (EC2, Lambda, S3), Docker, PostgreSQL, Redis

Tools: Cursor, Lovable, GitHub Actions, Pytest

EXPERIENCE

Software Engineering Intern (AI/ML)

AutoUnify

May 2025 – Aug 2025

San Francisco, CA

- Built model scoring microservice (FastAPI, PostgreSQL) on Vertex AI tracking accuracy, p50/p95 latency, and cost per 1K tokens across production LLM endpoints.
- Used **Cursor (Claude)** for API design, test generation, and code review, accelerating iteration while maintaining production quality.
- Implemented CI gates (pytest, lint, JSON Schema) with regression thresholds; authored reusable prompt packs reducing manual corrections by **25%**.

ML Research Team Member — Amazon Nova Trusted AI

Virginia Tech (Team HokieTokie)

Jan 2025 – Jun 2025

Blacksburg, VA

- Placed **1st** (Tournament 2) and **2nd** (Tournament 1); built evaluator components reducing jailbreak success by **46%**.
- Built data pipelines generating adversarial prompts across **100+ CWE** and **83 MITRE ATT&CK** behaviors with CI-backed analysis.
- Co-designed evaluator APIs and documentation enabling fast, reproducible extensions by other research teams.

Intelligent Process Automation Intern

Colgate-Palmolive GBS

Jan 2024 – Jun 2024

Mumbai, India

- Built **3 production chatbots** (HR, IT, procurement) with FastAPI, automating **70%** of inquiries and saving **200+ hours/month** across 400+ employees.
- Reduced response time from 30s to 10s using Redis caching and async I/O on **1K+ daily interactions**.
- Built monitoring dashboards with incident runbooks, reducing mean time to resolution for operations team.

PROJECTS

LunaFlow — Cycle-Aware Planning Platform

[GitHub] [Demo]

React, Node.js, Python/FastAPI, Docker, Google Gemini, ElevenLabs

- Built planning system in **24 hours** with 4 integrated views (calendar, Kanban, brainstorm, vision board); implemented bidirectional Google Calendar/Tasks sync via OAuth 2.0 eliminating duplicate data entry across platforms.
- Architected multi-service system (React, Node.js gateway, Python/FastAPI AI) with Google Gemini for task extraction from voice (ElevenLabs) and natural language symptom tracking; deployed at **lunaflow.work** with CI/CD, health checks, and automated rollback.

CareRoute — Multi-Agent Medical Tourism Orchestrator

[GitHub]

FastAPI, PostgreSQL, A2A SDK, Docker, AWS EC2

- Built orchestration system in **36 hours** at Codefest 2025 (4th Place) coordinating **6 agents** (flight monitoring, voice calls, notifications, hotel, hospital, accessibility) via JSON-RPC 2.0; successfully processed **12 end-to-end test bookings** with real voice calls and emails.
- Implemented workflow automation (flight ETA triggers), retry logic, HIPAA-compliant audit logging, and APScheduler for background jobs; deployed to AWS EC2 with PostgreSQL and Flutter frontend; used **Lovable and Cursor** for rapid development.

ACHIEVEMENTS & AWARDS

- **Amazon Nova Trusted AI Competition:** 1st Place (Tournament 2), 2nd Place (Tournament 1)
- **Codefest 2025 (Marriott):** 4th Place + Honorable Mention