

Harsh Sangani

+91 8511808931 | harshsangani15@gmail.com | [Portfolio](#) | [LinkedIn](#) | [Github](#) | [Codeforces](#)

OVERVIEW

AI/ML Engineer focused on building the systems around AI models. Experienced in developing production-grade AI applications, retrieval systems, agentic workflows, and multimodal AI solutions that improve decision-making and automate complex operations. Passionate about solving problems end-to-end, from product thinking and system design to deployment, evaluation, and optimization.

EDUCATION

Pandit Deendayal Energy University
B.Tech in Computer Science

July 2022 – June 2026
Gujarat, India

EXPERIENCE

AI Engineering Intern
The Product Folks

Jan 2026 – Present
Bangaluru, India

Talent Hub · Semantic Talent Search Platform | *ChromaDB, GPT-5-mini, Streamlit, Python*

[Live Demo](#)

- Built a hybrid vector and keyword search platform indexing **25,000+ candidate profiles** on 1024-dim mxbai embeddings, cutting recruiter shortlisting from 3+ hours to under 30 seconds (~99% faster).
- Designed a two-stage LLM pipeline (intent parse, then top-100 re-rank) lifting **top-10 relevance by ~60%** over pure vector search across 200+ test queries.

Content Velocity Suite · Reels + pSEO | *Claude Cowork, Higgsfield MCP, Webflow MCP, n8n*

- Composed a UGC pipeline auto-publishing **5+ Instagram reels/week** with competitor-engagement signals scraped via n8n, cutting content cycle time by ~70%.
- Launched a programmatic SEO engine shipping **10+ Webflow blog pages/week** with contextual CTAs to Jobs and Courses (LinkedIn job postings scraped through a parallel n8n workflow).

Lead Outreach Pipeline · WhatsApp Automation | *Selenium, Google Sheets API, Telegram Bot, OpenAI*

[GitHub](#)

- Delivered a single-sender 3-touch outreach engine with anti-ban guardrails (20 msgs/day, 36-hour follow-ups) and AI reply classification (Positive/Negative/Neutral) **automating 100% of cold outreach** for the recruiting team.

Software Developer Intern
The Special Character

May 2025 – July 2025
Gujarat, India

Court Booking Backend | *Node.js, Payload CMS, Medusa.js, OpenAI*

[GitHub](#)

- Developed **20+ REST endpoints and 10+ Payload CMS collections** powering slot booking, payments and self-serve owner dashboards across 100+ courts (~40% drop in venue ops overhead); added an LLM-powered query parser converting free-form text into structured DB filters.

PROJECTS

Product Teardown Studio | *FastAPI, Next.js 14, Playwright, Pydantic, OpenRouter*

[GitHub](#)

- Built a two-stage multimodal LLM pipeline (Gemma-3-27B vision plus GPT-OSS-120B analyst) producing **16-section structured teardowns and a 10-dimension radar scorecard in ~15 seconds** (vs. 2–6 hour manual analyst baseline), with a 3-tier JSON-repair waterfall hitting **98%+ completion** and an MD5-keyed 14-day cache cutting repeat latency to under 100 ms.

TeleMedGen | *Gemini 2.5 Flash, CNN, Node.js, MongoDB, Socket.IO, WebRTC*

[GitHub](#)

- Developed a patient & doctor telemedicine portal with **CNN-based dermatology triage (3–5 s/image)** and Gemini chat replies (1.5–2.5 s), AI conversation summaries cutting doctor chart-review by ~60%, WebRTC video consultations and multilingual STT/TTS across 4 Indian languages.

Shree Lalji Fabric Store | *Node.js, Express, MongoDB, React*

[Live Demo](#)

- Shipped a full-stack store with **15+ secure REST APIs** and Excel-based bulk import handling 1000+ SKUs in under 60 seconds (~70% manual onboarding effort cut), plus templated PDF catalog generation reducing catalog turnaround from hours to minutes.

TECHNICAL SKILLS

Languages: Python, JavaScript, TypeScript, Java, C, SQL, HTML, CSS

Frameworks: FastAPI, Next.js, React.js, Node.js, Express.js, Tailwind CSS, Spring Boot, Pydantic, Socket.IO, WebRTC, Selenium, Playwright, BeautifulSoup

AI / ML & Data: LLMs (GPT-5/4, Gemini 2.5, Claude, Gemma), LangChain, LangGraph, RAG Pipelines, Prompt Engineering, Vector Embeddings, ChromaDB, Pinecone, Semantic Search, CNN, MCP (Model Context Protocol), Multimodal AI, Agentic Workflows

Developer Tools: Git, GitHub, Docker, Postman, n8n, Claude Cowork, Higgsfield MCP, Webflow MCP, MongoDB, PostgreSQL, MySQL, Redis, OpenRouter, AWS (basics)

CERTIFICATIONS

IBM SkillsBuild – Cybersecurity Fundamentals | Python 101 for Data Science (PY0101EN) – IBM Cognitive Class

LEADERSHIP & COMMUNITY

Competitive Programmer, *Codeforces* (2023–25): Specialist rating 1492, Global Rank 624 in Round-1105 across 14,000+ participants.

Hackathon Participant, *Adobe & Odoo* (2024–25): built end-to-end MVPs in 24/36-hour sprints.