

Saivamshi Jilla

AI Engineer | Generative AI & Agentic Systems

Mumbai, India | saivamshijilla04@gmail.com | +91 89833 55654 | LinkedIn | GitHub | Portfolio

Summary

AI Engineer specializing in Large Language Models, Retrieval-Augmented Generation (RAG), and agentic automation systems. Independently owned and shipped a production multilingual Generative AI evaluation platform end-to-end (Gemini, multi-engine OCR, 80+ node n8n orchestration), achieving 85%+ accuracy across 500+ verdicts. Skilled in LangChain/LangGraph agent orchestration, vector search, and full-stack delivery – turning ambiguous AI problems into deployed, production-ready systems.

Technical Skills

Generative AI & LLMs: LLMs, RAG, Prompt Engineering, LangChain, LangGraph, Multi-Agent Orchestration, OpenAI API, Anthropic (Claude) API, Hugging Face Transformers, NLP, Computer Vision, Reinforcement Learning

Vector Search & Data: ChromaDB, FAISS, Weaviate, Qdrant, PostgreSQL, MySQL, MongoDB, Supabase

Automation & Cloud: n8n Automation, AWS, GCP, Docker, FastAPI, Flask, REST APIs

Languages & Web: Python, C++, JavaScript, TypeScript, React, HTML/CSS

Tools: Git, Postman, Selenium, Playwright, Cursor, Claude Code, Figma, Canva

Experience

Smiligners

Mumbai, India

AI Engineer Intern

Apr 2026 – Aug 2026

- Independently owned and delivered a production multilingual (Hindi/English/Marathi) AI answer-sheet evaluation system end-to-end, using Gemini 3.5 with multi-engine OCR (Mathpix + Google Vision), achieving 85%+ accuracy across 500+ verdicts.
- Designed a self-learning feedback loop that captured evaluator corrections to continuously improve scoring accuracy over iterations, backed by guardrails and automatic fallback across OCR engines with confidence-based routing.
- Orchestrated an 80+ node n8n workflow that cut the manual-review rate to 17%, and built idempotent MySQL schemas with Python ingest services for audit-complete, traceable evaluation storage powering client dashboards.
- Developed AI-integrated features for standalone AdTech products, including LLM-powered content generation, intelligent targeting logic, and automated campaign-optimization workflows.

LinCXO

Navi Mumbai, India

AI Engineer Intern

Dec 2025 – Mar 2026

- Built an LLM-powered resume and job-description parsing, matching, and ranking system achieving 85%+ accuracy, reducing manual candidate-screening time by 60%.
- Designed Python and Playwright web-scraping pipelines to aggregate and normalize job listings at scale; built LLM agents with function calling to automate internal hiring workflows, including a prototype AI video-generation agent.

Projects

AgentBay | *AI Agent Marketplace* – Live Link

July 2026 – Present

- Architected a full-stack AI agent marketplace using a scalable monorepo, enabling integration and independent deployment of 10+ specialized agents (code, research, writing, data analysis) with a secure sandbox execution environment.
- Designed agent orchestration supporting dynamic tool-calling, inter-agent communication, and stateful session management using LangGraph and FastAPI, with a React UI improving agent discoverability and onboarding.

Food Package & Freshness Detection | *Computer Vision* – GitHub

Nov 2024 – Dec 2024

- Engineered an end-to-end CV pipeline combining YOLOv8 for object detection and LLaMA for contextual analysis, achieving 90% food-quality assessment accuracy on a custom dataset of 5,000+ images.
- Automated extraction of product metadata (brand names, ingredient lists, expiry dates), reducing manual quality-inspection time; built a dynamic UI supporting live video streams and batch image uploads.

Autonomous War-Torn Navigation System | *Robotics & AI* – GitHub

Oct 2023 – Feb 2024

- Developed a drone-based autonomous rescue system achieving 80% object-detection accuracy with optimized A* path-planning for dynamic obstacle avoidance; engineered a low-latency Bluetooth protocol for real-time path transmission and sensor fusion across LIDAR and camera inputs.

Leadership & Achievements

- Led a team of 30+ student engineers as Technical Director of the institution's Robotics Team.
- 1st Place – National Level Product Showcase for Embedded Systems and Microcontrollers.
- Runner-Up – Institute-level Web Development Competition; Semi-finalist – International E-Yantra Robotics Challenge (EYRC).

Education

Vidyavardhini's College of Engineering and Technology

Mumbai, India

Bachelor of Engineering in Computer Engineering, Mumbai University – CGPA: 7.6/10

May 2026

Certifications

- NVIDIA Deep Learning
- Python Fundamentals