

Education

Marathwada Mitra Mandal's College of Engineering
B.E. in Information Technology, GPA: 9.01/10

Pune, Maharashtra
Dec 2021 – Jun 2025

- **Patent Published:** AI-Powered Exam Assessment System using OCR-free multimodal LLMs (2025)

Professional Experience

NeuraMonks
AI Engineer

Ahmedabad, Gujarat
Oct 2025 – Present

- Built and deployed **Agentic AI systems** using **CrewAI, Langraph, and Google ADK**, automating multi-agent workflows and boosting throughput by **35%**.
- Optimized LLM pipelines with dynamic context pruning and memory reuse, reducing **token cost by 40%** and latency by **25%**.
- Designed a modular **retrieval-reasoning-planning API stack** and integrated multimodal validation, improving accuracy by **20%**.

Kontiki Innovation Labs
GenAI Full Stack Developer

Remote / Pune, Maharashtra
Aug 2025 – Sept 2025

- Played a key role in bringing **Jarwiz.ai** from development to production, ensuring scalability, performance, and feature completeness.
- Refactored and optimized backend APIs, improving code quality and reducing system latency by **30%**.
- Developed **developer productivity tools** that accelerated internal development speed by **4×**.
- Worked with **Next.js, FastAPI, AWS (RDS, EC2, S3), Vercel, Pinecone**, and LLM APIs.

Rocket.new
AI Engineer

Surat, Gujarat
Jan 2025 – Jul 2025

- Shipped **Rocket.new** by engineering GenAI features and reducing operational costs by **40%**.
- Automated workflows via internal AI tools, improving team velocity by **30%**.
- Prototyped LLM + Vector DB features using **Gemini, Claude, GPT**.

Technical Skills

Programming:	Python, C++, Java, SQL
Machine Learning:	Scikit-learn, XGBoost, LightGBM, MLflow, SHAP
Deep Learning:	TensorFlow, PyTorch, Keras, HuggingFace Transformers
Generative AI:	Agno AI, CrewAI, Langraph, Google ADK, LangChain, Prompt Engineering
NLP / Vision:	NLTK, spaCy, OpenCV, ESRGAN, VGG16, YOLO
MLOps:	Docker, AWS (SageMaker, EC2), Azure ML, Git, FastAPI, Jupyter
Databases:	MySQL, PostgreSQL, SQLite, MongoDB
Vector DBs:	Pinecone, Chroma, Milvus, Qdrant
Web Dev:	Flask, Django, ReactJS, NodeJS, ExpressJS, HTML/CSS/JS

Projects

- **Browser Automation Extension (GenAI, Automation):** Built an LLM-driven Chrome extension using FastAPI and Gemini 2.5 to automate multi-step browser actions from natural language. Designed a client-server feedback loop for autonomous task execution until completion.
- **Frontend to Backend Code Generator (GenAI, Systems):** Created a GPT-4.1 + Claude 3.7 powered system to analyze frontend repos and auto-generate production-ready Node.js APIs with DB schema, auth, and Postman tests. Integrated heuristic parsing to prioritize endpoints and optimize deployment flow.
- **AI-Powered Exam Assessment System (Patent, Multimodal LLM):** Developed an OCR-free system for evaluating handwritten answer sheets and diagrams, achieving **85%+ human-grade accuracy**. Incorporated ViT-based diagram classification; patent published in 2025.
- **Coding Assistant CLI (GenAI, ADK):** Designed a CLI coding assistant with Gemini 2.5 and Pinecone memory for multi-turn problem solving. Capable of reasoning over local codebases and suggesting fixes.
- **GuidoAI: AI Career Guidance Chatbot (NLP):** Implemented a RAG-based chatbot in Flask with MongoDB backend to suggest optimal career paths. Engaged 1.2K+ students with a **93% satisfaction rate**.