

Yashovardhan Tiwari

+91-7073215996 | yt282003@gmail.com | linkedin.com/in/yashovardhan-tiwari | github.com/codezeewrangler

EDUCATION

Vellore Institute of Technology

Bachelor of Technology in Computer Science and Engineering; CGPA: 8.08/10.0

Bhopal, MP

Expected May 2026

PROJECTS

ComponentForge | Python, Claude API, LangChain, Vite, React, Angular

- Built an agentic pipeline that converts natural language prompts into production-ready Angular and React components, featuring an autonomous **self-correction loop** with up to 3 LLM retry attempts per generation.
- Engineered a rule-based validator agent that enforces design-system compliance – catching rogue hex colors, border-radius mismatches, and syntax errors before output, with errors fed back into the LLM for auto-remediation.
- Implemented multi-turn interactive REPL mode with full conversation history, enabling iterative component refinement through natural follow-up instructions (e.g., “make the button rounded”).
- Integrated a Vite-powered live preview app that hot-reloads the generated component in the browser instantly on each generation, reducing visual feedback time to near-zero.

InsightATS Resume Builder | Python, FastAPI, LangChain, React, Docker, FAISS

[Live Demo](#)

- Engineered a full-stack ATS resume analysis platform that parses resumes and job descriptions through a LangChain LLM pipeline, computing structured keyword alignment scores between candidate profiles and job requirements.
- Built a multi-agent workflow with dedicated agents for skill extraction, gap analysis, and suggestion generation – transforming unstructured resume text into actionable improvement reports.
- Implemented semantic similarity scoring using Hugging Face sentence embeddings and FAISS vector search, enabling context-aware skill matching beyond exact keyword lookup.
- Deployed FastAPI backend with Docker on Render, integrating Redis response caching to achieve consistent **sub-300ms** API latency under standard load; secured with JWT authentication and per-endpoint rate limiting.

Nexus AI Agent | Python, LangChain, FastAPI, Redis, Hugging Face

- Integrated LangChain tool calling and persistent contextual memory, enabling the agent to autonomously decompose complex user queries into sequential, dependency-aware sub-tasks.
- Optimized inference throughput via Redis-backed response caching and prompt batching, reducing redundant LLM calls on repeated or semantically similar queries and achieving a **2x** speed improvement on cache hits.
- Engineered an async task queue with automatic retry logic and exponential backoff, improving resilience against API timeouts and transient model provider failures.
- Secured all endpoints with JWT authentication, rate limiting, and CORS protection; implemented structured JSON logging for end-to-end request traceability in production.

TECHNICAL SKILLS

Languages & Scripting: Python, Java, SQL, Bash

Backend Frameworks: FastAPI, LangChain, LangGraph, Node.js

AI / LLM: RAG Pipelines, LLM Orchestration, Agentic Workflows, Prompt Engineering, Hugging Face Transformers

Vector Search & Storage: FAISS, ChromaDB, Embedding Pipelines

Databases & Caching: Redis, PostgreSQL, MongoDB

DevOps & Cloud: Docker, AWS, Render, CI/CD

CERTIFICATIONS

OCI AI Foundations – Oracle (Sept 2025) | Cloud AI Infrastructure, Model Deployment

Generative AI Professional – Oracle (Oct 2025) | LLM Integration, Prompt Engineering

Machine Learning Specialization – Coursera, Andrew Ng / Stanford (Dec 2023)