Vaijayanti Deshmukh

AI Engineer | Forward Deployed Engineer | LLM, RAG & Agentic Systems
Washington, DC 20037 | (202) 913-9075 | deshmukhvaijayanti5@gmail.com
linkedin.com/in/vaijayanti-deshmukh | github.com/VJGit1 | vaijayantideshmukh.vercel.app

Summary

AI Engineer with an M.S. in Data Science (GPA 3.84) and 2+ years of experience developing Python backend services, LLM applications, and RAG pipelines. Experienced in agentic workflows, evidence validation, guardrails, and human-in-the-loop review. Skilled in deploying applications with FastAPI, Docker, Redis, and cloud services, and collaborating directly with executives to translate ambiguous requirements into practical AI solutions.

Professional Experience

Respondr Inc. McLean, VA
AI Developer Intern *Jun 2025 – Sep 2025*

- Built AI application features using Python/Flask, React, TypeScript, PostgreSQL, JWT authentication, and WebSockets, integrating LLM services into operational workflows.
- Designed a RAG pipeline using OpenAI embeddings and semantic search to match emergency response plans with 95% accuracy, reducing retrieval time from minutes to seconds.
- Developed GPT-4 chat assistants and RAG workflows for incident triage, reducing manual triage time by 40%.
- Implemented severity scoring and decision-tree guardrails backed by PostgreSQL to structure recommendations and support human review.
- Partnered with the CEO and CTO to turn business requirements into an AI roadmap and deployable solutions.

Cognizant Ltd. Pune, India
Programmer Analyst *Nov 2022 – Jul 2024*

- Improved reporting performance by 16% through SQL query optimization and data retrieval enhancements.
- Collaborated with cross-functional teams in Agile environments across software development and testing workflows, supporting reliable integration of business requirements into production systems

Projects

Deal Memo Auto Generator | Python, FastAPI, RAG, Pydantic, Docker, Redis, Gemini Embeddings

- Architected an end-to-end, 8-step RAG and agentic workflow spanning document ingestion, semantic chunking, embeddings, synthesis, fact-checking, evidence validation, and structured output.
- Built evidence-gated generation, discrepancy detection, confidence scoring, and human review to produce reliable, traceable LLM outputs.
- Automated conversion of due diligence documents into investment memos and CRM-ready JSON, cutting drafting from hours to minutes.
- Deployed Docker Compose and Redis/RQ workers for long document jobs, reporting progress and failures via FastAPI.

CondoGPT: Agentic Real Estate Intelligence Platform | LangGraph, LangChain, RAG, FAISS, PostgreSQL, Flask

- Architected a LangGraph ReAct agent using GPT-4o-mini and LangChain to translate natural-language requests into validated SQL queries for real-estate analysis.
- Built a FAISS/OpenAI RAG pipeline for semantic retrieval, entity matching, and address resolution over large property datasets.
- Integrated structured data, vector search, conversational workflows, and APIs into a full-stack app with location-aware insights.

Credit Risk Modeling and Bias Detection (Capstone) | Python, XGBoost, SHAP, FastAPI Demo

- Built an ML system combining fairness auditing and SHAP explainability, achieving a 1.04 Disparate Impact Ratio and 0.70 Macro F1 across 51,000+ loan applications.
- Built a FastAPI inference engine with a versioned model registry and a profit-vs-fairness dashboard (73.5% approval rate).

Bonsai Biome: Emotion-Driven Journal | React, TypeSafe Jev, Notion API, Vercel Live Demo

- Built and deployed a journaling app that scores each sentence with five-way mood probabilities and grows a canvas tree per day; a rate-limited Vercel serverless proxy keeps API keys server-side.

Education

George Washington University Washington, DC
M.S. in Data Science, GPA 3.84 *Aug 2024 – May 2026*
Global Leaders Award (merit-based, academic excellence) | VP of Finance, GW Net Impact

Savitribai Phule Pune University Pune, India
B.E. in Computer Engineering (Minor: Data Science), GPA 3.76 *Aug 2018 – May 2022*

Certifications

Microsoft Azure Fundamentals (AZ-900) | Microsoft AI Fundamentals (AI-900) | Microsoft Power Platform Fundamentals (PL-900)

Technical Skills

Languages: Python, SQL, C++, TypeScript, JavaScript
Generative AI / LLM: RAG, Agentic AI, LLM Orchestration, n8n, Tool Calling, Prompt Engineering, LangChain, LangGraph, GPT-4, Claude, Gemini, MCP, Embeddings, LLM Evaluation, Evidence Validation, Guardrails, Human-in-the-Loop Review
AI / Machine Learning: XGBoost, Scikit-learn, SHAP, NLP, Bias Detection, Model Risk, Panda, Numpy
Data / Retrieval: Vector Search, FAISS, Semantic Search, PostgreSQL, MongoDB, Supabase, Neo4j, Pandas, SQL Optimization
Backend / Cloud / Tools: FastAPI, Flask, REST APIs, Pydantic, Docker, Redis, AWS, Azure Functions, Azure AI Search, CI/CD, Git