

Vaishnavi Jaiswal

AI / ML ENGINEER

CONTACT

- ✉ vaishnavijaiswal.work@gmail.com
- ☎ +91 8353945347
- 📍 Lucknow, India
- 🌐 vaishnavijaiswal.dev
- 🐙 github.com/vaishnavijaiswalwork-hub
- 🔗 linkedin.com/in/vaishnavi-jaiswal

AI / LLM

- RAG Architecture
- LangChain
- LangGraph
- OpenAI API
- Gemini API
- Qdrant
- Vector Search
- Prompt Engineering
- AI Agents
- mem0 · Neo4j
- FastAPI
- Ollama

PROGRAMMING

- Python
- JavaScript
- React
- Flutter / Dart
- HTML / CSS
- Java

DEV TOOLS

- Git / GitHub
- Firebase
- Vercel
- Figma
- REST APIs
- Linux

EDUCATION

B.Tech — Computer Science
Uttaranchal University · Dehradun
Expected 2029

Class XII — Science
Lucknow Public School · 81.2%
2024 – 2025

Class X
Lucknow Public School · 87.8%
2022 – 2023

LANGUAGES

English Fluent
Hindi Fluent

PROFESSIONAL SUMMARY

AI/ML Engineer with production experience building LLM systems, RAG pipelines, and autonomous agents. Shipped AI-powered features on platforms serving 200K+ monthly active users. Proficient across the full AI engineering stack — prompt engineering, vector databases, retrieval optimization, and LLM integration using OpenAI, Gemini, LangChain, and local models.

PROFESSIONAL EXPERIENCE

AI & Product Engineer

DIFMO PVT. LTD.

📅 Jan 2026 – Jun 2026 · Lucknow, India

- PressHop** 55K+ users · 200K+ monthly active · 19 cities (UK, USA, India)
 - Built domain-grounded RAG systems for content discovery; grounded LLM responses in PressHop's knowledge base, reducing hallucination on live product queries
 - Shipped AI-assisted product features end-to-end — prompt engineering, vector integration, QA, and deployment on a live scaling platform
 - Led technical troubleshooting, regression testing, bug triage, and build validation during rapid feature releases
 - Owned full product lifecycle — requirements gathering → architecture → implementation → stakeholder reporting

ADDITIONAL CLIENT DEPLOYMENTS (VIA DIFMO)

Sajjad Husain Law Associates

LegalTech RAG platform — case law, judgments, multilingual legal discovery. 50K+ monthly visitors, 10K+ legal professionals.

Algorithm Infrastructure

Enterprise RAG assistant over 1,200+ delivered-project knowledge base covering cloud, cybersecurity, networking, and telecom documentation.

Tylax Construction Chemicals

Designed and deployed a responsive product website covering 6 categories: waterproofing, adhesives, admixtures, grout, decoratives, and tile & stone care.

FEATURED AI / LLM PROJECTS

RAG Document Q&A Pipeline

PDF ingestion → chunking → OpenAI embeddings → Qdrant vector search for grounded, sourced Q&A

Python · LangChain · OpenAI · Qdrant

Tool-Calling AI Agent

Autonomous plan→action→observe loop; selects and executes tools with strict structured JSON output

Python · OpenAI · Pydantic

Memory-Aware Chatbot

Persistent conversational memory with multi-turn context via mem0 + Neo4j + Qdrant

mem0 · Neo4j · Qdrant · OpenAI

Stateful LLM Agent

Graph-based query routing with MongoDB-backed conversation state and multi-step orchestration

LangGraph · LangChain · MongoDB

Local LLM API Server

FastAPI wrapper for self-hosted Ollama/Gemma 3 enabling private on-device LLM inference

FastAPI · Ollama · Gemma 3

Code Journey — Web App

Multi-user learning tracker with Firebase, PIN-separated profiles, persistent progress. Live in production.

JavaScript · Firebase · Vercel