

Gaurav Kumar

gaurav.12bhindwar@gmail.com | +91 9006045930
github.com/gauravbhindwar | linkedin.com/in/gauravbhindwar

Summary

GenAI Engineer specializing in building production-grade AI systems using LLMs, RAG pipelines, and vector databases. Experienced in retrieval optimization, embeddings, and scalable backend systems using FastAPI, Redis, and Docker.

Technical Skills

GenAI & LLMs: OpenAI API, Prompt Engineering, Function Calling, RAG
Retrieval Systems: LangChain, Pinecone, Embeddings, Chunking, Semantic Search
Backend: FastAPI, Node.js, REST APIs, RBAC, Microservices
Frontend: Next.js, React, Tailwind CSS
Databases: MongoDB, PostgreSQL, Vector DBs
Systems: Redis, MQTT, Event-Driven Architecture
DevOps: Docker, CI/CD, GCP, Linux
Languages: Python, JavaScript, C/C++

Experience

GenAI Engineer Intern

Dec 2025 – Present

Simplify Consulting

- Designed and deployed **RAG-based AI systems** using vector databases.
- Built scalable **FastAPI services integrating LLM APIs**.
- Optimized latency using **Redis caching and retrieval optimization**.
- Developed event-driven workflows using MQTT and Pub/Sub.

Software Developer Intern

Jun 2025 – Sept 2025

AGILITYAI Pvt Ltd

- Built scalable applications using Next.js, Node.js, MongoDB.
- Migrated backend to **microservices architecture** and automated CI/CD.

Full Stack Developer (MERN)

Jan 2025 – Jun 2025

Software Development Center, Manipal University Jaipur

- Led development of **MentorLink** serving 500+ users with RBAC.
- Improved API performance via query optimization and caching.

Projects

Autonomous RAG System – Financial Intelligence (MoneyControl)

- Built automated **web-crawling RAG pipeline** using Puppeteer.
- Designed **hourly auto-refresh pipeline** for embedding updates.
- Implemented semantic retrieval (chunking + top-k search).
- Reduced hallucinations via prompt engineering.

AI RAG Chatbot – Knowledge Base

- Built end-to-end RAG pipeline with embedding-based retrieval.
- Improved accuracy via chunk tuning and retrieval optimization.

MentorLink – Mentorship Platform

- Built platform serving **500+ users** with RBAC and dashboards.

Education

Manipal University Jaipur

B.Tech in Computer Science

2026

GPA: 7.91

Certifications

- Full Stack Web Developer (Udemy), Next.js Development (Meta Brains)
- DSA in C/C++, CCNA Enterprise Networking (Cisco), Introduction to AI (IBM)

Achievements

- 2nd Rank – International Project Expo
- Top Performer – AGILITYAI Internship