Pulkit Sharma

Generative AI Engineer | Backend & Automation Specialist

✉ pulkit2693@gmail.com | 📞 +91 9354672119 | 💼 LinkedIn | GitHub

Summary

Generative AI Engineer with expertise in **LLMs, backend integration, and chatbot automation**. Skilled in **Python, Java, TensorFlow, and Spring Boot**, with experience building **AI-powered assistants, computer vision models, and intelligent Q&A systems**. Delivered projects such as **Vakta-Chatbot, CodeSage, and GenAI manga panel generator**. Hands-on with **Gemini API, Ollama, and Mistral API** for real-world conversational AI applications. Recognized in **hackathons** and passionate about transforming research into impactful AI solutions.

Education

Bachelor of Technology – Computer Science & Engineering (Specialization in AI & ML) Oct 2022 – Present
Dronacharya College of Engineering, Gurugram
CGPA: 8.14 / 10 (Till 5th Semester)

Skills

- **Programming**: Python, Java, C, SQL
- **AI/ML**: TensorFlow, Keras, Scikit-learn, NumPy, Pandas, Stable Diffusion
- **LLM APIs & Automation**: Gemini API, Mistral API, Ollama, n8n automation
- **Databases & Visualization**: MySQL, PostgreSQL, Power BI, Tableau, Excel, Matplotlib, Seaborn
- **Backend & Tools**: Spring Boot, Git, GitHub, VS Code, IntelliJ IDEA, Postman, Linux, Jupyter Notebook

Work Experience

Gen AI Intern – Liberin Technologies Pvt. Ltd., Noida Aug 2025 – Present
- Built **n8n-powered chatbots & AI assistants** by integrating LLMs (Ollama, Mistral, Gemini).
- Integrated **Gemini API** into conversational systems for improved contextual responses.
- Developed backend services with **Spring Boot** for AI-driven applications.
- Designed **prompt-engineered solutions** for domain-specific AI assistants.

Summer Intern – Computer Vision & Deep Learning, IGDTUW, New Delhi Jun 2024 – Aug 2024
- Created **YOLOv8 brain tumor detection** model with **30% faster inference**.
- Built a **CNN-based facial expression recognition** model with 59.2% accuracy.
- Contributed to deployment-ready research pipelines.

Projects

Vakta-Chatbot – Multi-Tenant RAG Chatbot Platform (Spring Boot + Ollama + Qdrant)
- Engineered a **full-stack, multi-tenant platform** in Java Spring Boot allowing users to create, manage, and deploy custom RAG chatbots.
- Implemented a **RESTful API for asynchronous URL ingestion**, storing embeddings in a dedicated **Qdrant vector collection per chatbot** to ensure strict data isolation.
- Integrated with **Ollama LLMs** for retrieval-augmented responses grounded in tenant-specific knowledge.
- Tech stack: Java Spring Boot, Ollama, Qdrant, REST APIs.

CodeSage – AI-Powered Development Assistant (Spring Boot + Ollama + PostgreSQL)
- Built a **full-stack AI development tool** offering **AI-powered code review, automated test generation, and refactoring suggestions**.
- Designed a **parallel file analysis engine** that generates vector embeddings for each source file, enabling semantic understanding of large codebases.
- Stored embeddings in **PostgreSQL with pgvector**, allowing semantic search and similarity queries across source code.
- Tech stack: Java Spring Boot, Spring AI, PostgreSQL (pgvector), Ollama.

GenManga AI – Generative AI Manga Panel Creator
- Developed a **Stable Diffusion-based system** to generate manga panels from text prompts, producing high-quality A4 pages.
- Engineered a **dynamic panel layout algorithm** for adaptive multi-frame composition with varying sizes.
- Integrated **speech bubble rendering & emotion-driven dialogues**, simulating natural storytelling.
- Optimized pipeline to generate a complete manga page in under **5 minutes**, improving usability.
- Tech stack: Python, Stable Diffusion, PIL, Custom Prompt Engineering.

Certifications

- Spring Boot for Java Developers — Udemy (Aug 2025)
- Complete Data Science, ML, DL, NLP Bootcamp — Udemy (Jul 2025)
- Data Science & Analytics — HP Foundation (Jul 2025)
- Mathematics for Machine Learning — Great Learning (Jul 2024)
- *Foundational*: Python Training, Java Training — IIT Bombay (Dec 2023), LearnWik (Apr–Jun 2023)

Achievements & Hackathons

- NCAT 2025 (Category 3): All India Rank 12 in Category 3 of national level aptitude test.
- ML Quest Hackathon: Flight Price Prediction secured a rank in Top 10.
- DataSphere (SpaceCon 2025): Finalist, led project as the team leader in the competition.
- Google ML Crashcourse: Earned 13 ML badges in google Machine Learning crashcourse.

Leadership & Responsibility

ACM Student Chapter, DCE
- Treasurer (2024 – Present): Managed finances, led events, hackathons.
- Core Web Team (2023 – 2024): Organized technical activities & chapter events.