

AKASH KATTE

AI & Generative AI Consultant | Cloud Deployment Expertise | Machine Learning Specialist

+91 9322192237 akashkatte316@gmail.com linkedin.com/in/akashkatte Jath, Maharashtra



SUMMARY

AI & Generative AI Engineer with hands-on experience designing, building, and deploying scalable LLM-powered solutions, including RAG systems, NL2SQL pipelines, and AI-driven agents. Skilled in Python, LangChain, TensorFlow, PyTorch, and cloud platforms (AWS/Azure/GCP). Experienced in vector databases, LLM evaluation, prompt engineering, and data preprocessing for enterprise-grade AI. Strengthened applications by integrating 5+ AI agents and deploying production-ready AI services. Eager to contribute to Deloitte's Digital Internal Audit team by developing AI-led risk, analytics, and automation solutions that enhance audit efficiency and decision-making.

EDUCATION

- Bachelor of Technology (B.Tech)
SKN Sinhgad College of Engineering - 9.62 CGPA
08/2022 - 06/2026 Korti, Pandharpur
- Higher Secondary Certificate (HSC)
M.V. High School and Junior College - 75.67%
06/2021 - 05/2022 Umadi, India
- Secondary School Certificate (SSC)
Bharat High School and Junior College - 88.80%
06/2019 - 05/2020 Utagi, India

EXPERIENCE

- Intern as Python Backend Developer
Rezoomex
08/2025 - 09/2025
Focused on Production level Backend development using python
 - Designed and deployed backend microservices using **Flask, FastAPI, Django** for production-level AI pipelines.
 - Built an end-to-end **RAG system** integrating Sentence Transformers + ChromaDB + LLM inference.
 - Automated data cleaning, preprocessing, and ETL workflows, reducing manual effort **40%**.
 - Implemented **REST APIs** for scalable AI and data operations.
- Trainee GenAI Engineer
Pathways to Parenthood
10/2025 - Present
Gives the solutions to who want to be a future parent.
 - Built and deployed **AI-driven agent systems** using LLMs (GPT-4.1, Gemini), improving software automation by **20%**.
 - Developed **custom GenAI pipelines** including retrieval, context management, and prompt engineering for real-world applications.
 - Integrated **5+ AI agents** into production systems, enhancing decision automation and workload efficiency.
 - Worked with **RAG architectures, embeddings, vector search**, and multi-step reasoning pipelines.
 - Collaborated with teams to translate business requirements into **LLM-driven automation specifications**.

SKILLS

Machine Learning

GenAI

Deep Learning

AI Tool Handling

NLP

Django

FLASK

RESTAPI

MYSQL

Python

Large Language Model

React

Node.js

MongoDB

Internal Audit

AWS

Azure

GCP

SQL

TensorFlow

PyTorch

Keras

Risk Management

Critical Thinking

Decision-Making

Continuous Learning

Interpersonal Skills

Leadership

Langchain

CERTIFICATION

- Introduction to Data Science
CISCO | 2025
- AWS
Intellipaat | 2025
- Python
HackerRank and Kaggle | 2025
- Machine Learning and Generative AI
Gradiuus | 2025
- 100 Days of Python Coding
TalentBattle | 2024
- Getting Started with Enterprise Data Science
IBM | 2023
- Power BI and Data Analytics
VOIS Edunet Foundation | 2023

PROJECTS

- GENAI Powered RAG Tutoring System
07/2025 - 08/2025
Developed a full-stack RAG application allowing natural language queries on CSV data. The Python/Flask backend uses Sentence Transformers for embeddings, Chroma DB for vector storage, and Google Gemini to generate accurate, source-attributed answers.
 - Designed full-stack **RAG architecture** using Sentence Transformers, Google Gemini, and ChromaDB.
 - Implemented **data preprocessing**, document chunking, embedding pipelines, and retrieval evaluation.
 - Built interpretable, audit-friendly responses with **source attribution**, improving trust and transparency.
- GENAI Powered NL2SQL system
07/2025
Built a Natural Language to SQL (NL2SQL) system using Python and Open AI GPT-4.1. Implemented a semantic search retriever with Sentence Transformers to provide the LLM with dynamic, context-aware few-shot examples, resulting in highly accurate SQL generation and step-by-step reasoning.
 - Built an LLM-driven **Natural Language → SQL** agent using GPT-4.1 with semantic search based few-shot examples.
 - Designed robust data pipeline with **retrievers, context engines, and validation checks**.
 - Improved query accuracy using **dynamic RAG-based prompt construction**.
 - Evaluated LLM responses using accuracy, consistency & execution success metrics.

TRAINING / COURSES

- Machine Learning and Generative AI
(Gradiuus | 03/2025 to 09/2025)
Completed intensive training in Machine Learning and Generative AI from Gradiuus, covering topics from foundational algorithms to advanced Large Language Models.