Atharv Kharmate

awishkar193@gmail.com | +919309203067 | Portfolio | LinkedIn | GitHub

Professional Summary

Results-driven **Software Engineer – GenAI & Data** with strong experience in **software development, backend systems, and AI-driven applications**. Proficient in designing and building **scalable APIs, data pipelines, and document intelligence platforms** that transform complex, unstructured data into actionable insights.

Education

•Bachelor of Engineering in Information Technology *2021-2025*
SCTR'S Pune Institute of Computer Technology, Pune CGPA: 8.09

•Higher Secondary Certificate (HSC) *2021*
Vynkatrao High School and Junior College, Ichalkaranji Percentage: 93.50%

•Secondary School Certificate (SSC) *2019*
G.D.K. Sainiki School, Tasgaon Percentage: 94.40%

Experience

•Software Engineer – GenAI & Data *Oct 2024 – Present*
Dataeaze Systems Pvt. Ltd. Pune, India

- Designed and developed **AI-driven document intelligence systems** for regulatory compliance and audit automation.
- Built an end-to-end **Risk & Control Matrix (RCM) generation pipeline** from SOP and regulatory PDFs using LLMs and semantic search.
- Implemented **regulatory impact analysis** by mapping regulations to SOP processes using vector embeddings and similarity scoring.
- Developed backend pipelines using **Python, LangChain, Azure OpenAI**, ensuring explainable and auditor-ready outputs.

•Data Engineer *May 2024 – Oct 2024*
Dataeaze Systems Pvt. Ltd. Pune, India

- Worked on **data ingestion and processing pipelines** for structured and unstructured enterprise data.
- Implemented data validation, transformation, and storage workflows to support downstream AI models.
- Assisted in building scalable backend services and improving data reliability for analytics use cases.

Projects

•IDAS ML Ingestion & Inference Pipeline for Optical Fiber Data *Sep 2024 - Ongoing*
Python, C++, Data Analysis & Visualization, NVIDIA Jetson

- Developed high-throughput ML ingestion and inference services for real-time optical fiber sensor data.
- Developed config-driven acquisition and processing pipelines, supporting dynamic runtime configuration updates.
- Added history-mode replay and robust service lifecycle management.

•Automated Risk & Control Matrix (RCM) Generation System *May 2024 - Dec 2024*
Python, LangChain, Azure OpenAI, React, PDF Processing

- Built an AI-powered system to automatically generate **Risk & Control Matrices** from SOP and regulatory PDF documents.
- Implemented dual-mode processing for SOPs and regulatory documents with clear marking of AI-assisted controls.
- Integrated semantic search and citation highlighting to ensure **audit traceability**.
- Reduced manual RCM preparation effort by up to **70–80%**.

•Regulatory Impact Analysis Automation *Aug 2024 - Dec 2024*
Python, Vector Embeddings, LLMs, Excel Automation

- Designed a system to analyze regulatory changes and identify impacted SOP processes automatically.
- Used vector similarity search to map regulations to relevant business processes.
- Generated auditor-friendly impact reports with High/Medium/Low classification and Excel export.

Technical Skills

Languages: Python, C++, Java, SQL, JavaScript
GenAI / ML: Prompting, LangChain, LangGraph, Azure OpenAI, LLMs, Embeddings, RAG
Backend: FastAPI, Data Pipelines, Document Processing
Databases: MySQL, MongoDB, VectorDB, Firebase
Frontend: HTML, CSS, ReactJS, NodeJS, ExpressJS
Tools: Git, GitHub, Postman, NVIDIA Jetson, VMs
Coursework: Object-Oriented Programming, DBMS, Data Structures & Algorithms, Artificial Intelligence & Machine Learning

Achievements

•Customer Hero: Appreciation received from Aditya Birla Capital Ltd. *2025*
•2nd Runner Up: PICT CoDay - Conducted by NICE Systems Co. *2024*
•Best Performance: Award for demonstrating excellence during internship *2023*