Atharv Kharmate

awishkar193@gmail.com +91-9309203067
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 *2021-25*
SCTR'S Pune Institute of Computer Technology, Pune CGPA: 8.09

•12th - Higher Secondary Certificate (HSC) *2021*
Vynkatrao Highschool and Junior Collage, Ichalkaranji Per: 93.50%

•10th - Secondary School Certificate (SSC) *2019*
G.D.K.Sainiki School, Tasgaon Per: 94.40%

Experience

•Software Engineer – GenAI & Data *Oct 2025 – 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 2025 – Oct 2025*
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 2025 - 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 2025 - Dec 2025*
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 2025 - Dec 2025*
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, Azure OpenAI, LLMs, Embeddings, Semantic Search
Backend: Fast APIs, Data Pipelines, Document Processing
Databases: MySQL, MongoDB, Vector Databases
Frontend: ReactJS
Tools: Github, Postman, Jetson
Coursework: OOPS, DBMS, DSA

Achievements

•Customer Hero : Appreciation received from the ABCL *2025*
•2nd Runner up : PICT CoDay - Conducted by NICE Systems Co. *2024*
•Best Performance : Award for demonstrating excellence during the internship. *2023*