

LALIT KUMAR

+91 8867400741 | lalitsheoran876@gmail.com | [Linkedin](#) | [GitHub](#)

SUMMARY

AI Engineer (3+ years) with a strong Python development foundation, building innovative solutions from scratch. At Phyllo, engineered a GenAI based CMO co-pilot that opened new revenue channels and accelerated growth.

EDUCATION

Dhirubhai Ambani Institute of Information and Communication Technology
MSc IT

Jul 2020 – May 2022
CGPA: 7.8

PROFESSIONAL EXPERIENCE

Machine Learning Engineer
Phyllo

Aug 2024 – Present

- Designed an **Multi-Agent RAG pipeline** for a Decision Intelligence tool for Brand Managers using **LangGraph** and reasoning models (GPT-o3, Claude Sonnet4), reducing market research and opportunity/threat analysis time from **1–2 weeks to under 10 minutes**.
- Built an **orchestrator** that breaks down and routes each user query to the appropriate agent—SQL generator, vector search, third-party API connectors, or web search agents.
- Architected a **hierarchical vector store** (Weaviate), logged directory-level context in MongoDB, and indexed it in Elasticsearch—enabling **95% accurate** semantic search and reducing query-routing latency by **60%**.
- Containerized **modular agents and orchestration logic** using Docker, deployed on AWS EC2, exposed via FastAPI endpoints, and scaled the system to handle **1,000+ req/min** at **99.9%** uptime with CI/CD workflows.

Data Scientist
Fractal Analytics

Jun 2022 – Jul 2024

- Designed a bullet ballistic image-matching pipeline using fine-tuned **CLIP Vision Transformer** and **ResNet-50** to identify if the same gun was used across crime scenes, accelerating forensic investigations and achieving accuracy up to **92%**.
- Built a agentic pipeline for mortgage-based loans by extracting data from rental/operating statements using fine-tuned **table-transformer** and OCR, then processing the data to determine loan eligibility and generate proformas—reducing manual effort by **70%** and processing time by **60%**.

TECHNICAL AND SOFT SKILLS

- Languages and Related Technologies:** Python, GoLang, FastAPI, Azure, AWS, PostgreSQL, MongoDB, Docker, Kubernetes, RabbitMQ, Redis, GIT, Github, CI/CD, Kibana, Datadog, Microservices, Selenium
- AI/ML & Automation :** NLP, Transformers, Fine-Tuning (LoRA), PEFT, Scikit-learn, Tensorflow, Pytorch, RAG, Prompt Engineering, LangGraph, LangChain, LlamaIndex, MCP, Vector Databases, DuckDB
- Soft Skills:** Problem-Solving, Leadership, Cross-team collaboration, Business communication

CERTIFICATIONS

- AI-102: Microsoft Certified Azure AI Engineer Associate
- AZ-204: Microsoft Certified Azure Developer Associate
- Fractal Certified Associate PySpark Developer
- Databricks Accredited Lakehouse Platform Fundamentals

PUBLICATIONS AND ACHIEVEMENTS

- Published research paper titled **Deep Learning Based Automated Localization of Anterior Commissure and Posterior Commissure Landmarks in 3D Space from 2D Three-Plane MRI Localizer Slices of the Brain**, accepted at **ICMLDE 2022**. ([Paper Link](#))
- Recognized for improving custom video category screening accuracy to **94%** (from 68%) using **VideoMAE + LLMs** hybrid pipeline
- Solo Skydiver** certified by the **United States Parachute Association (USPA)** with over 100 jumps