

# Omolaiye Victor Kola — Machine Learning & AI Engineer

Email: victorkolaomolaiye@gmail.com | Location: Abuja, Nigeria | +2348106723379 | [LinkedIn](#) | [GitHub](#)

---

## PROFESSIONAL SUMMARY

I am a dynamic Machine Learning and AI Engineer with expertise in building and deploying production ready ML and GenAI systems, including RAG pipelines, intelligent agents, and MLOps workflows. Proficient in traditional ML (TensorFlow, YOLOv8, JAX) and GenAI (LangChain, embeddings, LLM orchestration), delivering scalable, secure, and observable solutions. Proven track record reducing manual effort by 40% and enhancing product relevance through semantic search, vector DBs, and CI/CD automation. Seeking junior or mid-level ML/GenAI/MLOps roles to drive innovative, data-driven solutions.

---

## Technical Skills

- **Programming & Data:** Python, SQL, Pandas, NumPy, Jupyter, OpenCV.
  - **Machine & Deep Learning:** Scikit learn, TensorFlow, JAX (Flax), XGBoost; YOLOv8; transfer learning, HPO, forecasting.
  - **Generative AI & NLP:** AI, Google Gemini, Hugging Face, LangChain, LLM fine-tuning, prompt engineering, chain-of-thought, embeddings, reranking.
  - **Retrieval & Vector DBs:** FAISS, Pinecone, Milvus, Chroma, semantic search, chunking.
  - **MLOps & Devops:** Docker, FastAPI, Flask, Gradio, Streamlit, Git, GitHub Actions, CI/CD, ONNX, monitoring (Prometheus/Grafana, Langsmith)
  - **Cloud & Storage:** Cloud & Storage: AWS (S3), Azure (AI Search, OpenAI, Foundry), DigitalOcean, PostgreSQL (local, neon, supabase), Redis, Cloudflare R2.
  - **Security & Practices:** RBAC/ABAC, secure ingestion, data privacy, performance profiling, system design.
- 

## PROFESSIONAL EXPERIENCE

### ML Engineer – African Institute for Artificial Intelligence (February 2025 – Present)

- Designed and developed a mentorship matching recommender system.
- Designed, developed and deployed a secure RAG Knowledge Assistant with Langchain, Pinecone, AWS S3, and PostgreSQL (neon), implementing RBAC/ABAC controls for role-specific document retrieval with zero data leakage. Reduced response time to seconds and improved retrieval precision.
- Built and managed and self hosted deployment options, integrated monitoring hooks (Prometheus/Grafana), and authored runbooks for incident handling and data retention.

### AI Engineer – Mensah Philosophic Circle (January 2025 – Present :contract role)

- Co-developed Hey Nova AI, an Azure OpenAI-powered educational agent with live internet access, using FastAPI, Azure AI Search, embeddings, and Redis for 200+ concurrent users. Added automated summarization and question-generation, aimed at focused study of course materials and boosting student engagement.
- Created Azure AI Foundry-based academic recommender system, designed to boost student engagement via personalized insights.
- Designed a YOLOv8/InceptionResNet-based premise security prototype (90%+ accuracy) integrated with FAISS for event retrieval, reducing loss incidents.

### AI/ML Engineer – Zummit Africa (January 2024 – June 2024)

- Development of a Zero-Shot Content Moderation Bot using Hugging Face LLMs (BERT Models) cutting manual review load by 40%. Built retraining pipelines and evaluation scripts for continuous model improvement.
- 

## SELECTED PROJECTS

- **Rag Assist:** Designed and developed a LLM powered multi tenant customer service solution using Gemini, Langchain, Python, superbase, pinecone, cloudflare R2 buckets that allows businesses to simply upload their documents and serve their customers a customer support solution instantly.
- **CV Ranker:** Built a FastAPI + LangChain app using Gemini embeddings to semantically rank CVs against job descriptions, generating CSV/PDF shortlists. Deployed on DigitalOcean with CI/CD and unit tests.
- **Research Assistant Tool:** Developed an LLM-powered tool to ingest research materials, auto-generate literature-review tables, and export to CSV/PDF, streamlining academic workflows.
- **TensorFlow-to-JAX Migration:** Ported ML models to JAX/Flax for performance optimization, demonstrating adaptability to modern frameworks.

---

## EDUCATION

B.Tech, Estate Management | Federal University of Technology, Minna (2024)

Relevant course work: Statistical Analysis, Econometrics , Data driven valuation, Calculus.

---

## CERTIFICATIONS

- ChatGPT Prompt Engineering for Developers - DeepLearning.AI (2025)
- REST APIs with Flask – Udemy (2024)
- AI/ML Specialization (Stanford Online): Advanced Learning Algorithms, NLP, Reinforcement Learning - Coursera (2023)
- Data Science Virtual Experience - Forage (British Airways, KPMG) (2023)

---

## ACHIEVEMENTS

- Architected RAG systems with sub-second retrieval and secure access for multi-tenant environments in production.
- Demo ready GenAI applications (Hey Nova AI, CV Ranker) with 60s walkthroughs for hiring outreach.
- Demo-ready: can present HeyNova.ai, Knowledge Assistant (managed and self-hosted options), CV Ranker or Research Assistant on request.
- Reduced content moderation workload experimentally by 40%.

---

## AVAILABILITY

Open to full-time, part time, or contract ML/GenAI/MLOps roles (remote-friendly except within Abuja). Demo-ready for Hey Nova AI, CV Ranker, and RAG Knowledge Assistant (managed/self-hosted), Rag Assist, etc upon request .