Abdulhakeem Salawu

AI Software Engineer | AI Engineer

hakeemsalawu3@gmail.com | *https://www.linkedin.com/in/abdulhakeem-salawu/*
35 Ogunfunmi Street, Ikotun, Lagos, 200654, Nigeria | *09036467962*

AI/ML Engineer specializing in Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning with expertise in Agentic AI, Large Language Models (LLMs), Generative AI, and Natural Language Processing (NLP). Proven track record designing and deploying LLM agents, Retrieval-Augmented Generation (RAG) pipelines, and ML-based automation to improve business process efficiency. Skilled in MLOps and LLMOps including CI/CD for ML, model monitoring, experiment tracking, model registry, and scalable data pipelines. Proficient with transformer-based architectures (LLaMA, GPT, Hugging Face), LangChain, LlamaIndex, embeddings, and vector databases (FAISS, Pinecone, Chroma). Experienced with predictive analytics, recommender systems, and computer vision solutions. Adept at building cloud-native AI systems on AWS, GCP, and Azure using Docker, Kubernetes, microservices, and serverless deployments, with strong focus on Responsible AI, security, and compliance (GDPR, HIPAA).

Work Experience

AI Software Engineer

May 2025 - Present

Novada | *Australia*

As an AI Software Engineer, I design, build, and deploy scalable software solutions powered by machine learning and artificial intelligence.
My role blends core software engineering with AI development, from writing production-grade code and integrating AI models into building intelligent systems to solve real-world problems.
I'm focused on creating robust, efficient systems that turn complex data into real-world impact.
I leverage AI to deliver innovative products and drive measurable business impact.

- Developed and deployed Agentic AI systems, integrating LLMOps, LLM Agents and RAG to automate business processes, resulting in a 40% increase in efficiency.
- Built and optimized Generative AI and NLP applications using transformer-based architectures (LLaMA, GPT, Hugging Face) with LangChain and LlamaIndex, integrating embeddings, knowledge graphs, and semantic search into cloud-based MLOps platforms (AWS Sagemaker, GCP Vertex AI, Azure ML). Implemented CI/CD, experiment tracking (MLFlow, Weights & Biases), and data version control (DVC), enabling 70% faster model deployment cycles and reducing operational overhead by 35%, while delivering scalable,
- Developed and deployed Agentic AI systems, LLM agents, and Retrieval-Augmented Generation (RAG) pipelines, integrating vector databases (FAISS, Chroma), reducing hallucinations in NLP applications and custom data pipelines to enhance contextual reasoning and information retrieval.
- Collaborated cross-functionally with product, engineering, and business stakeholders to align AI capabilities with business goals, ensuring measurable impact.
- Implemented LLMOps and MLOps best practices—covering model monitoring, evaluation, model registry, CI/CD for ML, experiment tracking (MLFlow, Weights & Biases), prompt/version management, data version control (DVC), drift detection, continuous monitoring, and multi-cloud deployment (AWS Sagemaker, GCP Vertex AI, Azure ML). These initiatives reduced model drift incidents by 40%, improved deployment reliability by 50%.
- Engineered scalable data pipelines (ETL/ELT) and ML-based automation for classification, information extraction, and natural language understanding, while optimizing LLM inference with quantization,

pruning, distillation, distributed training, batching, GPU acceleration (CUDA, ONNX Runtime), and secure, compliant enterprise integrations.
- Designed and deployed Agentic AI systems, LLM agents, and Generative AI applications leveraging LangChain, LlamaIndex, Hugging Face, and transformer-based architectures (LLaMA, GPT) to power workflow automation, multi-step reasoning, and real-time decision-making. These solutions automated over 60% of repetitive business processes, and reduced task completion times by 45%.

AI Engineer & Project Manager Team Lead

Dec 2024 - Present

MyMogulMedia | *United States*

Managing cross-functional teams, designing and implementing automation processes, workflows & Chatbots to optimize repetitive tasks, saving a ton of operational time.
Developing, deploying, and maintaining AI-powered solutions using Python, TensorFlow, and PyTorch, with a Strong background in computer vision and NLP.

- Applied MLOps principles—including CI/CD for ML, model packaging, and deployment with Docker/Kubernetes—to ensure production readiness of AI solutions, resulting in a 50% reduction in deployment time
- Leveraged deep learning frameworks (TensorFlow, PyTorch) for model fine-tuning, transfer learning, and domain-specific adaptation of LLaMA and GPT models, achieving a 25% improvement in model accuracy and a 40% reduction in training time
- Designed and deployed production-ready AI systems on cloud platforms (AWS Sagemaker, GCP Vertex AI, Azure ML) using Docker/Kubernetes, API orchestration (FastAPI, Flask), and microservices architectures. Ensured compliance with GDPR and HIPAA while implementing responsible AI practices, resulting in 99.9% system uptime.
- Utilized Scikit-learn, Pandas, NumPy, TensorFlow, and PyTorch for experimentation, version control with Git/GitHub/GitLab, and cloud infrastructure via Terraform (Infrastructure-as-Code).
- Designed and deployed services with REST APIs, FastAPI, Flask, and GraphQL, using containerization (Docker), Kubernetes operators, microservices architecture, and serverless deployment for scalable AI integration.
- Applied prompt engineering, fine-tuning, instruction tuning, and parameter-efficient techniques (LoRA, PEFT) to improve LLM performance in zero-shot and few-shot learning scenarios.

AI Developer & Software Engineer

Dec 2020 - Present

Fiverr & Upwork | *Isreal*

Contributed to AI product development and strategy, applying business process automation with LLM agents to improve efficiency and decision-making.

- Optimized performance of ML models through quantization, pruning, distillation, distributed training, and HPC-based inference optimization, reducing latency and infrastructure costs while enabling real-time decision-making.
- Applied Responsible AI and Ethical AI practices with focus on data privacy, GDPR, and HIPAA compliance, while contributing to AI product development, strategy, predictive analytics, recommender systems, and business process automation to align AI initiatives with measurable business outcomes.
- Leveraged Scikit-learn, Pandas, and NumPy to design and test ML models during experimentation and feature engineering.
- Developed and deployed REST/GraphQL APIs using FastAPI and Flask, within a microservices architecture and serverless deployments for scalable AI integrations.

Core Skills

Core AI/ML:
Artificial Intelligence (AI), Machine Learning (ML), Deep Learning, Natural Language Processing (NLP), Computer Vision, Generative AI, Predictive Analytics, Recommender Systems

LLMs & Agents:
Agentic AI, LLM Agents, Retrieval-Augmented Generation (RAG), Prompt Engineering, Instruction Tuning, Fine-Tuning (LoRA, PEFT), Embeddings, Zero-shot / Few-shot Learning, RLHF (Reinforcement Learning with Human Feedback)

Category name:

MLOps / LLMOps:
CI/CD for ML, Data Version Control (DVC), Feature Store, Model Registry, Drift Detection & Monitoring, Experiment Tracking (MLFlow, Weights & Biases), Pipeline Orchestration (Airflow, Prefect, Kubeflow)

Infrastructure & Deployment:
API Development (REST, GraphQL, FastAPI, Flask, Django), Microservices Architecture, Containerization (Docker), Kubernetes, Kubeflow, Ray Serve, Serverless & Edge AI, Cloud Platforms (AWS Sagemaker, GCP Vertex AI, Azure ML)

Data Engineering:
ETL/ELT Pipelines, Data Cleaning & Transformation, Data Orchestration, Pandas, NumPy, Scikit-learn, SQL / NoSQL Databases, Vector Databases (FAISS, Pinecone, Chroma)

Tools & DevOps: Git / GitHub / GitLab, Terraform (IaC), Docker Compose, Linux/Bash, Jupyter, VS Code

Responsible AI & Compliance: Responsible AI, Ethical AI, Data Privacy & Security, GDPR, HIPA

Additional skills:
JavaScript, React, Component-based UI frameworks, Front-end architecture best practices (performance, security, usability), Product and design lifecycle collaboration, Writing well-tested code, Building high-fidelity in-app chat experiences, Crypto exchange experience, Onchain experience (Ethereum addresses, ENS, dApps, blockchain services), Customer-focused product ownership, Revenue-driving product development, Positive energy in team collaboration, Passion for crypto and blockchain, Thriving in high-pressure/mission-driven environments, Actively seeking feedback for improvement

Education

Ahmadu Bello University, Zaria.

Oct 2014 - Nov 2018

Bachelor of Science Computer Science

GPA: 4.33