

Olalekan A. OYELEYE

Hull, UK | +44 7453 795225 | olalekanoyeleye@yahoo.com | [LinkedIn](#)
[GitHub](#)

SUMMARY

Machine Learning Engineer with 3+ years of experience delivering end-to-end AI solutions across NLP, computer vision, and deep learning. Expert in Python programming with strong hands-on experience in building and deploying models using frameworks such as TensorFlow, PyTorch, Hugging Face Transformers, LangChain, and retrieval-augmented generation (RAG). Skilled in fine-tuning LLMs with parameter-efficient methods such as LoRA and QLoRA for domain-specific tasks, including summarization, named entity recognition(NER), and document classification. Proven ability to design auditable, modular ML pipelines that integrate structured and unstructured data, with a focus on performance, reproducibility, and operational reliability. Passionate about applied research and continuously exploring innovative approaches to solving real-world problems.

TECHNICAL SKILLS

[CLICK HERE FOR MY VISUALIZATIONS](#)

Programming & Scripting: Python, R, SQL, JavaScript, Bash

Machine Learning Frameworks: NumPy, Pandas, Scikit-learn, TensorFlow, Keras, PyTorch, Spacy, YOLOv8/v11&v12, SAM/SAM2,

MobileSAM, PaddlePaddle, LLMs (GPT-4, LLaMA, Mistral, DeepSeek etc.) Hugging Face Transformers, Transfer Learning, Fine-tuning (LoRA, QLoRA, PEFT), Retrieval-Augmented Generation (RAG), Prompt Engineering & Alignment, Topic Modeling & Text Classification, Named Entity Recognition (NER)

Data Processing & Analysis: Pandas, NumPy, Spark, Feature Engineering, Statistical Testing, Power BI, Tableau, Matplotlib, Seaborn, Plotly

Cloud & APIs: AWS Glue, Lambda, EMR, S3, Athena, Redshift, DynamoDB, Kinesis, IAM, SageMaker, SNS, EC2, CloudFormation

Docker, CI/CD, Infrastructure as Code, Model Deployment

Web Scraping & Automation: BeautifulSoup, Scrapy, Python Scripting

Model Optimization & Automation: Bayesian Optimization, Randomized Search, Hyperparameter Tuning, Python Scripting, FastAPI, Web Scraping (BeautifulSoup, Scrapy)

Data Visualization & Analysis: Matplotlib, Seaborn, Plotly, Power BI, Tableau

PROJECTS

Platelet Cell Classification with YOLOv8 & YOLOv11

Applied YOLOv8 and YOLOv11 for detecting and classifying platelet cells under varying biological conditions. Used image preprocessing, augmentation, and regularization techniques to improve model robustness. Achieved up to 89% accuracy with ensemble models, demonstrating the effectiveness of YOLO-based approaches in biomedical image analysis.

RAG-Based Conversational System for Web Document Retrieval

Built a lightweight Retrieval-Augmented Generation (RAG) app using LangChain, ChromaDB, and GPT-4o-mini for querying web-scraped content. Integrated BeautifulSoup for scraping, with a Gradio chatbot interface for interactive, context-aware responses.

LangChain Agent for Data Analysis

Built a LangChain-powered agent that leverages natural language processing to analyze structured datasets (e.g., Kaggle CSV files). The system uses a Pandas DataFrame agent connected to OpenAI's LLMs to parse, summarize, and answer natural language queries directly from CSV data. Designed for non-technical users, it enables intuitive, conversational data exploration and insight generation. Integrated capabilities for data visualization and summarization to enhance the accessibility and interpretability of complex datasets.

Map and Text extraction using SAM and MobileSAM

This project performs high-resolution instance segmentation and reference detection from scanned maps using, Segment Anything Models (SAM) and Mobile Sam from Meta AI(Object detection and Instance Segmentation), PaddlePaddle(Text Extractor) and Geospatial tools like GeoPandas and Shapely to save the shapes. The pipeline automates the detection of plot boundaries and their associated reference labels, exporting the final results as geospatial data.

NLP - Structured Text Classification with LLMs

Developed a robust text processing pipeline to extract and organize key information(Named Entity Recognition) from unstructured documents into a structured format making use of Large Language Models.

Automated Transcription and Meeting Minutes Generation using Hugging Face Models

Developed a pipeline leveraging Hugging Face speech-to-text models to transcribe multi-speaker audio recordings. Applied NLP techniques for speaker diarization, summarization, and topic extraction to convert raw transcripts into structured meeting minutes. Enabled automated generation of concise, action-oriented summaries suitable for internal documentation and reporting.

EDUCATION

University of Hull

2024

MSc Artificial Intelligence and Data Science(Distinction)

Dissertation Title – “Deep Learning Models for the Automatic Classification of Platelet Cells”

Achievements

- Conducted in-comprehensive research, applying a range of convolutional neural network (CNN) architectures for the automatic classification of platelet cells, focusing on performance optimization and diagnostic accuracy.
- Employed hyperparameter tuning techniques including Randomized Search and Bayesian Optimization to systematically improve model performance.
- Enhanced model effectiveness through transfer learning by combining multiple pre-trained models for more robust feature extraction.
- Demonstrated significant potential for clinical impact by automating platelet cell analysis, contributing to faster and more reliable diagnostic workflows.
- Research Paper (In Progress): Exploring the use of YOLOv8 and YOLOv11 for the classification of platelet cells, focusing on enhancing accuracy and efficiency in medical image analysis.

Obafemi Awolowo University

2016

MSc Economics

EXPERIENCE

ILT Networks (Remote)

Data Scientist

March 2022 – Present

- Integrated Retrieval-Augmented Generation (RAG) systems using LangChain and OpenAI's GPT-4o-mini, automating data retrieval from large datasets and enhancing information access.
- Developed and deployed a LangChain-powered agent to automate complex data queries, enabling non-technical users.
- Adapted pretrained models to domain-specific tasks through transfer learning and hyperparameter optimization, improving detection precision.

The Logistics Institute, University of Hull

Rail Data Administrator

March 2023 – June 2023

- Validated and digitized rail freight datasets for the NR+ platform
- Contributed to tools aimed at decarbonizing UK freight transportation
- Gained practical skills in transport data systems and geospatial processing

Delegation of European Union to Nigeria and ECOWAS

Economics and Trade Analyst(Assistant) - Applied Data Science & Automation

March 2020 – March 2023

- Applied data science techniques (data wrangling, feature engineering, hypothesis testing) to analyze regional trade data and inform EU-Africa trade policies and compliance with WTO/AfCFTA frameworks.
- Built and deployed Python automation pipelines to extract, process, and distribute economic indicators and trade alerts, reducing reporting time by 70%.
- Designed and maintained structured trade databases using SQL and pandas, improving data accessibility for cross-functional teams.
- Created custom dashboards and data visualizations (Matplotlib, Seaborn, Power BI) to support diplomatic reporting and briefings.
- Collaborated with EU policy units to deliver data-driven insights for high-level decision-making, including for the EU-Nigeria Business Forum 2022.

High Commission of Sri Lanka, Abuja

Data & Research Analyst

December 2016 – March 2020

- Conducted economic and trade research to support diplomatic strategy, using data science methods to analyze bilateral market trends and investment opportunities.
- Used data visualization to communicate insights, contributing to a 28% export increase in 2018.

- Maintained and managed internal datasets related to trade, visa processing, and consular services; streamlined data workflows to support mission operations.
- Supported public diplomacy through data-informed event planning and communication for initiatives like Sri Lanka National Day and 150th Ceylon Tea Anniversary.

CERTIFICATIONS/TRAININGS

AWS Cloud Quest Certificates:

- [Solutions Architect](#)
- [Cloud Practitioner](#)
- [Data Analytics](#)
- [Machine Learning](#)

Hands-on AWS Projects:

- Serverless ETL Pipelines (Glue, Lambda, S3)
- Big Data Processing on EMR (Hadoop, Spark)
- S3 Data Lake with Glue Catalog
- Real-time & Batch Ingestion (Kinesis, Lambda)
- Automated Pipelines (Step Functions, CloudWatch)
- SageMaker ML Pipeline (Training, Deployment)
- Textract & Comprehend for Unstructured Data
- IoT Data Pipelines (IoT Core, DynamoDB)
- Redshift Data Warehouse Design

LLM Engineering & Agentic AI – Udemy, 2025

Referees are available on request.