

JOY OLUSANYA

+234 09159945712 | joynaomi81@gmail.com |

 [LinkedIn](#) |  [joynaomi81](#)
 [@Naomi_joy](#)  [Portfolio](#)

CAREER SUMMARY

I am a passionate AI Developer and Machine Learning Engineer who loves exploring the potential of Natural Language Processing and Large Language Models. I design, develop, and deploy innovative AI solutions that tackle real-world challenges by blending state-of-the-art machine learning techniques with modern NLP tools. Proficient in Python and Julia, I employ cutting-edge frameworks to build scalable models and optimize data pipelines for peak performance. Beyond the technical side, I thrive on mentoring teams, managing projects with a collaborative spirit, and driving forward research initiatives that keep me excited about the future of AI and ML.

EXPERIENCE

• Atom Group

AI/LLM Developer

April 2025

Lagos, Nigeria

- Contributed to the development of a commercial AI agent, integrating Instagram scraping pipelines for enriched social media intelligence.
- Built an automated data collection system using Python and RapidAPI, enabling real-time scraping of posts, user profiles, and comments.
- Designed a modular, object-oriented scraping framework, improving reusability and reducing maintenance overhead.
- Developed a shortcode-to-media ID converter and integrated external APIs to enhance insight precision by 85%.
- Reduced scraping failure rates by 70% through robust exception handling and reusable utility functions.

FriendnPal

AI Developer

January 2023 - March 2025

Lagos, Nigeria

- Contributed to the development of AI models for a mental health chatbot using NLP, improving response relevance and patient engagement by over 80 percent.
- Developed Retrieval-Augmented Generation (RAG) pipeline, increasing the chatbot's contextual understanding and information retrieval accuracy by 75 percent.
- Implemented preprocessing and normalization systems for QA datasets, reducing training noise and improving model performance by percent.
- Conducted data transformation and analysis to validate model outputs, enhancing user experience and overall system reliability.
- Transcribed and curated French-English data to support machine translation model improvements, boosting multilingual response capability.

Rendo AI

ML/NLP Engineer

June 2024 - December 2024

Nigeria

- Designed a Python-based data extraction pipeline using **Requests**, **BeautifulSoup**, and **Selenium**, improving data acquisition efficiency by 45 percent across multiple web sources.
- Developed custom web scraping tools to automate the collection of structured data from diverse websites, enhancing dataset completeness and variety.

Brilla AI Project, Kwame AI

AI Developer

January 2024 - March 2025

Ghana

- Developed and trained deep learning-based speech-to-text models for multilingual ASR, with a focus on Ghanaian accent.
- Validated and curated diverse datasets, improving transcription accuracy and model robustness to phonetic variation and background noise.
- Built reproducible evaluation workflows to optimize model performance across accents and domains.
- Collaborated with data engineers to align preprocessing pipelines with training workflows, enhancing scalability.

Upwork

Freelancer

February 2024 - December 2024

Remote

- Data annotator: Transcribed and identified medical terms for building a speech recognition model.
- Analyzed data trends to inform model training and improve accuracy.

LEADERSHIP EXPERIENCE

• Deep Learning Co-Lead

February 2025 - Present

Data Science Nigeria Obafemi Awolowo University

- * **Event Organization:** Co-organized and executed deep learning events, collaborating closely with the lead to ensure the success and smooth operation of each session.
- * **Mentoring & Tutoring:** Provided mentorship and tutoring to team members and participants, Encouraging learning and growth in deep learning techniques and concepts.
- * **Bootcamp Management:** Played an important role in the successful management of deep learning bootcamps, ensuring high-quality content delivery, engagement, and participant success.
- * **Collaborative Leadership:** Worked alongside the lead to drive team initiatives, contributing to the development of educational materials and resources for both beginner and advanced learners in deep learning.

• AI/ Machine Learning Lead

December 2024 - Present

Women in Data Science Ile-Ife, Osun State, Nigeria.

- Led workshops and training sessions on AI and machine learning for women, promoting data-driven problem-solving skills.
- Partnered with chapter leads to design technical learning paths, create educational resources, and coordinate speaker sessions for beginner and advanced learners.

PROJECTS

• English–French Machine Translation Evaluation

June 2025

Tools: Python, Hugging Face Transformers (MBart50TokenizerFast, MBartForConditionalGeneration), BLEU, ROUGE, METEOR [GitHub](#)

- Evaluated the **facebook/mbart-large-50-many-to-many-mmt** model for English–French machine translation.
- Implemented **batch translations** using MBart50TokenizerFast and MBartForConditionalGeneration.
- Assessed translation quality using standard metrics: **BLEU**, **ROUGE**, and **METEOR**.

• AskMyPDF – Conversational Document Assistant

July 2025

Tools: Python, Streamlit, LangChain, Google Gemini API, FAISS, PyPDF2

[GitHub](#) | [Demo](#)

- Developed an AI-powered app that enables users to ask questions from uploaded PDF files in natural language.
- Integrated Google Gemini via LangChain and used FAISS for vector-based document retrieval (RAG pipeline).
- Designed a clean Streamlit chat interface with multi-question support and automatic response rendering.
- Enhanced document comprehension for students, researchers, and professionals.

• Automatic Essay Scoring (AES)

July 2025

Tools: Python, Transformers (DistilBERT), Pandas, Matplotlib, Seaborn

[GitHub](#)

- Built an automated essay scoring system using **DistilBERT-base-uncased**, fine-tuned on a public essay dataset.
- Conducted thorough **data cleaning** and **text preprocessing** to ensure high-quality model input.
- Applied **data visualization** techniques (Seaborn, Matplotlib) to analyze essay lengths and score distributions.
- Evaluated model performance using **RMSE** and **Pearson correlation**, achieving reliable scoring consistency.

• Yoruba Proverb Detection

June 2024

Tools: Python, NumPy, Pandas, Scikit-learn, NLTK, SpaCy, Transformers, TensorFlow

- Built a system to classify Yoruba proverbs using machine learning and transformers.
- Applied Naive Bayes, BERT, and AfroLM Active Learning to improve classification accuracy.
- Preprocessed text using NLTK for tokenization, lemmatization, and feature extraction.
- Fine-tuned BERT and AfroLM to capture Yoruba semantic and contextual features.

• Fake News Detection

June 2024

Tools: Python, NumPy, Pandas, Scikit-learn, NLTK, TF-IDF, Matplotlib, Logistic Regression, Decision Tree, Streamlit

[GitHub](#) | [Demo](#)

- Built a fake news classifier using Logistic Regression and Decision Tree models.
- Used TF-IDF and NLTK to extract features from large-scale news datasets.
- Achieved 98.5% accuracy with Logistic Regression; 99.5% with Decision Tree.

- Evaluated results using precision, recall, and confusion matrix plots.

- **Emotion Recognition in Speech**

Tools: Python, Librosa, NumPy, Pandas, TensorFlow, Keras, Scikit-learn, Matplotlib, LSTM

December 2024

[GitHub](#) | [Demo](#)

- Built an LSTM-based model to classify emotions from speech audio.
- Extracted MFCCs, chroma, and spectral contrast using Librosa.
- Normalized and preprocessed speech data for training.
- Achieved high accuracy in detecting emotions like anger, happiness, and sadness.
- Deployed model via Streamlit for real-time emotion recognition.

- **Breast Cancer Prediction**

Tools: Python, NumPy, Pandas, Scikit-learn, TensorFlow, Keras, Matplotlib

Feb 2025

[GitHub](#)

- Built a deep learning classifier to predict tumor malignancy with 95%+ accuracy.
- Handled missing values and performed feature selection for better accuracy.
- Visualized model results and feature importance using Matplotlib.
- Designed a pipeline ready for diagnostic system integration.

- **Toxic Comments Detection**

Tools: Python, NumPy, Pandas, Scikit-learn, NLTK, SpaCy, Matplotlib

Aug 2024

[GitHub](#)

- Developed an NLP model to classify online comments as toxic or non-toxic.
- Applied TF-IDF, tokenization, lemmatization, and stopword removal.
- Trained a Naive Bayes model with 89.9% accuracy.
- Evaluated performance using F1-score, precision, and recall.

EDUCATION

Obafemi Awolowo University

Bachelor of Arts in Linguistics and African Languages

Ile-Ife, Nigeria

Jan 2021 – July 2025

SKILLS

- **Programming Languages:** Python, Julia
- **Web Technologies:** Streamlit, FastAPI, Flask
- **Database Systems:** PostgreSQL, MySQL
- **Vector Databases & APIs:** Pinecone, Chroma, Tavily API, OpenAI API, Gemini API
- **LLM Application Frameworks:** LangChain, LangGraph, Retrieval-Augmented Generation (RAG)
- **Data Science & Machine Learning:** Pandas, NumPy, Matplotlib, Seaborn, Tableau, Power BI, Scikit-learn, TensorFlow, Keras, PyTorch
- **Deep Learning Frameworks:** CNNs, RNNs, LSTMs, Autoencoders, JAX
- **Natural Language Processing (NLP):** Scikit-learn, NLTK, SpaCy, BeautifulSoup, Hugging Face Transformers, Language Modeling, LLMs
- **Software Engineering:** Docker, Postman
- **DevOps & Version Control:** Git, GitHub, GitLab
- **Linguistics Software:** Praat, ELAN, FLEx, Audacity
- **Research Skills:** Literature Review, Experimental Design, Data Collection, Annotation, Dataset Curation, Model Evaluation, Statistical Analysis, Problem Solving, Language Documentation

POSTERS AND PUBLICATIONS

- [P.1] Olusanya Joy and Daud Abolade. (2025). **ÒWE-YOR: Leveraging Transformer-Based Models for Yoruba Proverb Classification**. Poster presented at *Deep Learning Indaba 2025*, Kigali, Rwanda.
- [T.1] Olusanya Joy. (2025). **Tone in Yoruba Automatic Speech Recognition: Evaluating the Impact of Tone Recognition on Transformer-Based ASR Models**. Undergraduate Thesis, *Obafemi Awolowo University*, Ile-Ife, Nigeria.

VOLUNTEER EXPERIENCE

Virtual Coach Supported participants with hands-on guidance through coding tasks. Enabled attendees to understand and complete their first web app. Skills: Mentorship, Technical Explanation, Python, Django	<i>Django Girls, May 2023</i>
Protocol Assistant Coordinated guest reception and ensured smooth event logistics during the campus event. Supported successful execution by managing protocol tasks and maintaining clear communication. Skills: Communication	<i>She Code Africa OAU, May 2024</i>
AI Mentor Taught Python and Machine Learning fundamentals through practical, beginner-friendly sessions. Contributed to the upskilling of over 50 participants. Skills: Mentorship, Leadership, Communication	<i>LevelUp Techies Bootcamp, Sep 2024 – Nov 2024</i>
Deep Learning/AI Tutor Taught deep learning fundamentals, including neural networks, CNNs, and model evaluation using TensorFlow and Keras. Empowered over 20 participants to build and deploy basic deep learning models through hands-on sessions and mentoring. Skills: Public Speaking, Curriculum Design, TensorFlow, Keras	<i>Data Science Nigeria OAU, Mar 2025 – Apr 2025</i>

PROFESSIONAL MEMBERSHIPS

• She Code Africa	<i>January 2020 - Present</i>
• Data Science Nigeria OAU	<i>January 2019 - Present</i>
• Masakhane NLP	<i>February 2023 - Present</i>
• Linguistics Island	<i>January 2024 - Present</i>
• AfriClimate AI	<i>June 2024 - Present</i>
• Cohere For AI	<i>May 2024 - Present</i>

CERTIFICATIONS

• Udemy: Natural Language Processing in Python for Beginners	2024
• Coursera: Data Science Fundamentals	2024
• Endangered Languages Projects (Webinar): Language Documentation Webinar	2024

ADDITIONAL INFORMATION

Languages: English Language, Yoruba Language, German Language, Nigerian Pidgin