BUSAYO AWOBADE

iambusayor.github.io ⋄ +234(0) 703 650 2407 ⋄ hawobade@gmail.com

Lagos, NG

RESEARCH INTEREST

Research Engineer with over 3+ years of experience in cutting-edge research and algorithm development to tackle real-world challenges. My research interests encompass low-resource NLP, speech processing, efficient machine learning, and multilinguality. I am seeking opportunities to engage in impactful research that addresses complex problems and advances the state and adoption of AI, particularly in low-resource communities worldwide.

EDUCATION

Bachelor of Science in Computer Science, Federal University of Agriculture, Abeokuta   2017 - 2023

Deeplearning.ai: Neural Networks and Deep Learning, Structuring Machine Learning Projects, Convolutional Neural Networks, Sequence Models;

PUBLICATIONS

- **Busayo Awobade**, ..., Tobi Olatunji. AfriVox: Probing Multilingual and Accent Robustness of Speech LLMs. *Under review @ ACL ARR 2025*.

- Gabrial Zencha Ashungafac, Mardhiyah Sanni, **Busayo Awobade**, Alex Gichamba, Tobi Olatunji. AfriSpeech-MultiBench: A Verticalized Multidomain Multicountry Benchmark Suite for African Accented English ASR. *IJCNLP-AACL 2025*.

- Chris Emezue, NaijaVoices Community, **Busayo Awobade**, et al. The NaijaVoices Dataset: Cultivating Large-Scale, High-Quality, Culturally-Rich Speech Data for African Languages. *InterSpeech 2025*.

- Mardiyyah Oduwole, ..., **Busayo Awobade**, et al. The Role of Data Augmentation in African Machine Translation.

- **Busayo Awobade**, Mardiyyah Oduwole, Steven Kolawole. What Happens When Small Is Made Smaller? Exploring the Impact of Compression on Small Data Pretrained Language Models. *AfricaNLP 2024*.

EXPERIENCE

Machine Learning Research | Intron   Nov. 2024 - Present

- Trained large-scale multilingual speech models for 11 African languages using a **novel curriculum learning** strategy.

- Built a robust pipeline for collecting and preprocessing spontaneous speech data, integrating **curriculum-based pseudo-labeling** and a **custom chunking algorithm**.

- Developed a **production-grade AfroTTS** system— a multispeaker, African-accented TTS platform supporting 40 distinct accents.

- Designed and implemented a **benchmarking framework** for evaluating multimodal and unimodal ASR models on African languages, with language-agnostic post-processing.

- Created the **Afrivox Benchmark**, a comprehensive ASR and AST evaluation suite covering **20 African languages**

Machine Learning Research | Babs   Jan. 2024 - Jun. 2024

- Trained and evaluated **TTS models** for 3 African languages (**Yoruba, Swahili, and Hausa**) using the **VITS architecture** on publicly available African speech corpus.

- Extended **Meta's MMS model to produce better natural sounding Yoruba and Hausa speech** using the common voice aligned corpus.

- Led the research on **speech-to-speech translation with expressivity transfer** which followed Meta's Seamless Expressivity work.
- Built a **cascaded speech-to-speech translation system** using ASR (**whisper**), TTT (**m2m100**) and expressive TTS model, **Bark**.

Machine Learning Researcher | MLCollective   Nov 2023 - Present

- Leading the research team for the local chapter (Nigeria); Helping young research enthusiast find their footing in ML research.
- Moderating talk sessions, meetups, presentations, and webinars on NLP.

Founding Machine Learning Engineer | ASAlytics (now NazariAI)   Mar. 2022 - Dec. 2022

- Awarded $**115k grant** in funding from the Algorand Foundation to build the **ASAlytics** platform;
- Set up **CI/CD pipeline** for model and application server using **GCP cloud build** and **GCP cloud run**.

Data Science, Intern | Zummit Info Labs   June. 2021 - Oct. 2021

- Built an object detection model to detect and classify road traffic signs using a **yolo** model coupled with a deep learning classifier to correctly classify detected traffic signs according to their respective classes for autonomous vehicles.
- Trained a race detection model for a machine learning system that detects racial profiling and racial bias amongst airport security using Google's **FaceNet** to generate face embeddings passed into a **Lightgbm** classifier.
- Used **Dlib**, **Mtcnn**, **OpenCV** to build face detection and recognition system for vehicle security.

PROJECTS & AWARDS

Algorand Foundation Grant Award - Algorand Foundation

- Awarded $**115k** to build an opinion mining platform for the algorand ecosystem. (2022)

Best Poster Awards - Deep Learning Indaba

- Won best poster awards at the 2023 Deep Learning Indaba for **What happens when small is made smaller?**.

Omdena Communities

- Led the machine learning team for the Columbia community to build a multilingual **bertTopic** model for both Spanish and English topic generation on tweets/reddit posts to monitor the 2021 Columbia nationwide project. (2021)
- Using **Yolo-v5**, **docker** and **AmazonEC2**; built a machine learning solution to speed up (**2x**) tea-leaf processing for the Sri-Lanka Tea industry. (2021)

DSN Data Science Hackathon - Data Scientist Network

- Finished **1st placed runner-up** amongst 750 skilled Data scientist at the DSN nationwide hackathon. (2020)

UmojaHack Africa - Umoja, Kenya

- Led my university to a **3rd runner-up finish as the 4th best AI university in Nigeria in 2021**.

AXA Nigeria 2020 Hackathon - AXA Mansard

- Led my university to a **5th place finish** at the AXA-Mansard Insurance claim hackathon. (2020)

VOLUNTEER ACTIVITIES

- Peer Reviewer AfricaNLP@ICLR '24, EMNLP '23, TinyPapers@ICLR '23.
- Co-organizer; Masakhane's EfficientNLP workshop at Deep Learning Indaba 2023.
- Co-Lead and Mentor at MLCollective's Nigeria local chapter.
- Ex-lead for the AI community at Federal University of Agriculture, Abeokuta;