Sharon Ibejih

Lagos, Nigeria.

📱 +2348128253909 • 🔗 https://sharonibejih.github.io/bio/

🔗 https://github.com/sharonibejih • ✉ sharonibejih@gmail.com

Education

BS(Hons), Physics

Michael Okpara University, Abia State, Nigeria. 2013 - 2017

CGPA: 4.06/5.00

Skills: ML — Software — Cloud

Stack: Python, SQL

Tools: NumPy, Pandas, Scikit-Learn, TensorFlow, PyTorch, MLflow, Streamlit, Flask, Docker, Kafka, Git, AWS

Work Experience

Insidatics AI Inc., Delaware – ML Software Engineer (Contract)

Feb. 2022 – Date

Insidatics is a machine learning startup with a product that helps financial institutions predict the credit worthiness of customer(s).

○ Tested the developed API endpoints using Postman and fixed detected bugs.

○ Developed Flask RESTplus API Documentation for the product.

○ Developed building materials price prediction model for a client.

○ Codebase maintenance.

RectLabs Inc., Lagos – AI Engineer

July 2022 – Sept. 2022

RectLabs is an AI startup that is building products such as a no-code/low-code AI platform for computer vision tasks and a synthetic data generation system.

○ Optical character recognition system for converting English and Yoruba written PDFs to TXT, leveraging pytesseract.

○ Deployed sports analysis and object auto-annotation AI models.

○ Researched and implemented image data generation, using Imagen.

Data Scientists Network (DSN), Lagos – Solution Analyst

Oct. 2020 – April 2022

○ Developed a Pidgin Speech Recognition model that gave the organisation an international contract for local Nigeria speech recognition in social media.

○ Experimented building a doctor's handwriting OCR model with KerasOCR and TrOCR for a research grant.

○ Co-wrote a proposal for the 2021/2022 Lacuna Funding for community-curated local dataset.

○ Developed an end-to-end audio analytics pipeline for an organisation to help them understand their call-center data. This entailed building a speaker diarizer (who spoke when) for a mono channel to separate the customer from the call agent. Then transcribe each speech chunk and analyse the output text. The success of this project, led to a bigger contract for the organisation.

○ Carried out machine learning research experiments for publications.

○ Co-wrote the AI Ethics document for the organisation.

○ Managed the development process of an EdTech product.

○ Led the product and solution briefings, ran live product demos, and prototypes.

Projects

NIGERIAN GOOGLE PLAYSTORE APPS ANALYTICS
- Collaborated with a team of four to scrape and analyse Google Playstore Nigerian apps using Python.
- The analysis involved EDA based on app genre and sentiment analysis of customer reviews on Playstore.
- Built product that gives real time sentiment analysis of any Nigerian Product, using their Playstore app ID.
- The product also provides a visual analytics of any selected app genre.
- Product **link**

NAMED ENTITY RECOGNITION IN PDFs
- Converted PDF pages to images and used an OCR system to convert the images to texts.
- Applied spaCy pretrained model to identify the named entities.
- Converted the output named entities to a DataFrame for display.
- Modularized and packaged the codes.

SPEECH RECOGNITION SYSTEM FOR NIGERIAN CHURCH SERMONS
- Coordinated the data collection and transcription of the Church's sermons.
- Developed a speech recognition model with the sermon data, experimenting with QuartzNet and wav2vec2 XLS-R pretrained models.
- Packaged the model trained on XLS-R for raw audio transcription and deployed to DockerHub.

Research Publications

Oyewusi, W. F., Adekanmbi, O., **Ibejih, S.**, Osakuade, O., Okoh, I., & Salami, M. (2021). AFRIFASHION1600: A Contemporary African Fashion Dataset for Computer Vision. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3968-3972).*

Oyewusi, W.F., Adekanmbi, O., Okoh, I., Onuigwe, V., Salami, M.I., Osakuade, O., **Ibejih, S.**. & Musa, U.A. (2021). NaijaNER: Comprehensive Named Entity Recognition for 5 Nigerian Languages. *arXiv preprint arXiv:2105.00810.*

Ibejih, S., Oyewusi, W. F., Adekanmbi, O., & Osakuade, O. (2022, March). EDUSTT: In-Domain Speech Recognition for Nigerian Accented Educational Contents in English. In 3rd Workshop on African Natural Language Processing

Community Activities

She Code Africa
DS/ML/AI Technical Coordinator   *Dec. 2020 - Jan. 2022*

AI Saturdays Lagos
Cohort 7 Co-organiser and Mentor   *Sept. 2021 - Jan. 2022*

ICLR 2022
ICLR Conference Volunteer   *April 2022*

AI Saturdays Lagos Meetup
Speaker: Opportunities in Machine Learning   *Sept. 2022*