

# Sharon Ibejih

Lagos, Nigeria.

📞 +2348128253909 • 🌐 <https://www.linkedin.com/in/sharonibejih/>

🌐 <https://github.com/sharonibejih> • ✉ [sharonibejih@gmail.com](mailto: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

---

### Data Scientists Network (DSN)

*Solution Analyst*

*Oct. 2020 – April 2022*

- Developed an in-domain Speech Recognition model for Nigerian Accented English Educational Contents, with a validation WER of 27%.
- Experimented building a doctor's handwriting OCR model with KerasOCR and TrOCR.
- Developed a Speaker Diarization system that separates multiple speakers recorded in a mono channel.
- Co-researched on AFRIFASHION dataset generation and trained baseline CNN fashion wares classification models.
- Developed a Nigerian Pidgin Speech-to-text model, with a word error rate of 33%.
- Led the team on the collection and annotation of Speech data.
- Carried out machine learning research experiments for publications.
- Drafted AI Ethics document for the organisation.
- Product Management for an EdTech product.
- Led product and solution briefings, ran live product demos, and prototypes.

### RectLabs Inc.

*AI Engineer*

*July 2022 – Sept. 2022*

An AI startup that is building products such as 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.
- Development and deployment of auto-annotation tool for object detection based on user-selected image classes.
- Research and implementation of image generation, using diffusion model.

### Insidatics AI

*Part-time ML Software Engineer*

*Feb. 2022 – Date*

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

- Testing API endpoints using Postman and bug fixtures.
- Developed Flask RESTplus API Documentation for the product using Swagger UI and deployed on Heroku.
- Developed and dockerized building materials price prediction model developed with XGBoost.

## 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](#)

### TEXT ANALYSIS OF ENDSARS TWEETS

- Scraped EndSARS tweets using SNScraper and preprocessed the data with Python.
- Adopted and evaluated Zero-Shot learning for emotion classification across seven classes.
- Keyword Extraction based on the emotion classes: Popular hashtags used during the EndSARS events were extracted from the tweets.
- Developed a sentiment analysis of the tweets. Logistic Regression model performed best using the CountVectorizer with an f1-score of 82%.

### 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., **Ibejeh, 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., **Ibejeh, S.** & Musa, U.A. (2021). NaijaNER: Comprehensive Named Entity Recognition for 5 Nigerian Languages. *arXiv preprint arXiv:2105.00810*.

**Ibejeh, 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