# Abib **Duut**

♀ 15 Potier Street, SE1 4UX, London.



### **EXPERIENCE**

### Present October 2022

#### Data Scientist | Laboratory Research Scientist, THE FRANCIS CRICK INSTITUTE, London

- > Developing computational methods for detecting dementia type diseases in the brain.
- > Compared the structure of responses of artificial networks to that of biological vision networks.
- > Created dashboards for neural data exploration, descriptive analysis & data visualization.
- > Fitting distributions & machine learning models to neural data.
- > Traced and labelled cells using Labkit and built a machine learning model in StarDist.

Python NVim Git Jupyter sklearn matplotlib emacs numpy opency scipy Zsh

#### July 2024 March 2024

# Data Scientist | Analyst, FREELANCE, Remote

- > Data imputation for the price of (agricultural) commodities for different locations in Accra.
- > co-developed an economic market model to simulate a market resource allocation and determines

Python nVim Git Zsh Lua

#### October 2022 December 2020

#### Data Scientist | Principal Research Assistant, Institutional Research, Ghana

- > Developed a model to infer the number of administrative staff for the university of Ghana.
- > Performed an attrition analysis which informed the upholding of student sports concession.
- > Served on a committee to assess the feasibility of implementing a document verification technology.
- > Improved the institute website, managed in Drupal.

Python Vi Git Pandas MT<sub>E</sub>X Bash

### December 2021 October 2020

#### ML/Data Engineer | Freelance, REмоте, Ghana

- > Developed parsers to check for data consistency and integrity.
- > Built scalable and reusable ETL pipelines for ingesting large data sets from multiple APIs.
- > Developed a model for vehicle attribute detection including make, model, and year.
- > Developed routines for data visualizations of large image datasets using Matplotlib.
- > Built Extract Transform Load pipelines for ingesting weather & energy data from APIs.
- > Built explainable models for the forecast of Electrical Energy demand from weather data.

Python PyTorch SQLAlchemy Pytest AWS

# July 2019 May 2019

# Machine Learning Engineer | Machine learning, PI CAMPUS, Rome

- > Worked on a real-time energy forecast and interpolation project (over 30M rows of data). Developed an RNN model with an annual savings benefit of 100 thousand euros.
- > Experimented with Sparse Gaussian models, Prophet, and some spectral methods.

Tensorflow AWS GPflow Pandas Jupyter Notebook FaceBook Prophet PyCharm

# September 2018 September 2017

### Scientific co-worker, FORSCHUNGZENTRUM JÜLICH, Germany

SMARTSTART 2 Fellow pour Computational Neuroscience

- > Co-developed an Argumentation Mining model, which won DL4NLP2018 award
- > Attempted developing a Generative Adversarial Network for Google Landmarks data.
- > Attempted using PredNet model to learn the latent representation of eye blinks to find an optimal strategy for blinking.
- > Explored the limits of state of the art Deep Convolutional Neural Networks including Inception, ResNet, VGG16, VGG19, by testing them on the task of Contour Integration.

Keras PyTorch Python Scikit-learn Docker Linux

## November 2015 February 2014

### Data Scientist | Associate Data Analyst, Ghana Meteorological Agency, Ghana

with travel to Kenya and Tanzania

- > Translated proprietary Genstat code to R code.
- > Developed a Markov Model for the prediction of amount and chance of rainfall.
- > Developed a reference report for each synoptic station in Ghana; this report gives an overview of weather data at each station including missingness.
- > Contributed to the development of ClimateObject, an R package specifically for the analysis of climate data in Africa; and the development of a Maproom for the Ghana Meteorological Agency sponsored by the International Research Institute for Climate and Society, Columbia University.

R Visual Basic GLM ATEX Sweave

# February 2014 January 2014

## Data Engineer | Data Assistant, RANCARD SOLUTIONS LIMITED, Ghana temporary

> Developed a dictionary in Excel for a huge data labelling task - a database of three million rows. Completed task in 3 weeks.

Microsoft Excel



# **EDUCATION**

- 2012 2013 African Institute for Mathematical Sciences, Ghana, MSc. in Mathematical Sciences
- 2007 2011 University of Ghana, BA. in Mathematics with Economics



# Training & Awards

- 2021 National Information Agency, Certificate, Tableau
- 2021 Coursera, Certificate, Cloud Big Data and Machine Learning Fundamentals
- 2021 **Coursera**, Certificate, Troubleshooting and Debugging Techniques
- 2021 Coursera, Certificate, Jenkins: Automating your delivery pipeline
- 2021 Coursera, Certificate, Building Batch Data Pipelines on GCP
- 2021 Coursera, Certificate, Introduction to Javascript: The Basics
- Coursera, Certificate, Deploying SAP on Google Cloud 2021
- 2020 Coursera, Certificate, Introduction to Git and Github
- 2019 Pi Campus, School of Artificial Intelligence, Certificate in Applied Machine Learning Full Scholarship, selected one of top two-percent of applicants
- 2017 Forschungzentrum Jülich, Germany one year SMARTSTART Fellowship Award in Computational Neuroscience
- African Institute for Mathematical Sciences (AIMS), Ghana, Master in Mathematical Science 2013 One year full scholarship
- 2005 2005 Best student for Physics, and Mathematics in high school after completing an accelerated (3-year) programme.



# **66** Referees

#### Professor Andreas Schaefer

Assistant Director

THE FRANCIS CRICK INSTITUTE, LONDON NEUROPHYSIOLOGY

andreas.schaefer@crick.ac.uk

NA

#### Professor Bernd J. Schroers

Co-Director of the Maxwell Institute CHAIR, ICMS MANAGEMENT COMMITTEE MACS, University of Edinburgh

@ b.j.schroers@.ac.uk

+ 44 (0) 131 451 8247