

Michael Dubem Igbomezie

Via Monte Rotondo, n.2, 67100, L'Aquila, Italy

☎ +39 350 103 0134 | ✉ michaeligbomezie@gmail.com | 📄 github.com/dub-em | 🔗 linkedin.com/in/michael-igbomezie-2901a5122

About Me

I am an Applied Data Scientist and Business Analyst, with strong inclination to AI and Deep Learning, Software Development and Big Data. I am quite passionate about improving the society and human rights with AI. Searching mostly for Data Science, Machine Learning and Data Analysis positions.

Specialties Natural Language Processing (BERT, LLM), Computer Vision, Reinforcement Learning, Deep Learning, Machine and Statistical Learning, Big Data (Apache Spark), Time Series (ARIMA), Database, TPL and API Development and Deployment, ML Lifecycle.

Education

Universita' degli Studi dell'Aquila

L'Aquila, Italy

MSc in Applied Data Science

Sep 2021 - Oct 2023

- Grade: 95%
- **Certificate:** [Link to Certificate](#)
- **Thesis Research:** [Link to Thesis Research](#)
- **Courses:** Advanced Techniques for Machine Learning, Statistical Theory and Methods, Big Data and Algorithm, Programming for Data Science, Networks and Decision Models, Open Data and Web Services, Business Intelligence and Analytics, Time Series Analysis and Forecasting

University of Port Harcourt

Rivers, Nigeria

BEng in Mechanical Engineering

Sep 2013 - Aug 2018

- Grade: 86%
- Designed Automobile Systems for HEPSSA program under Dr. Ojapah Mohammed
- Designed a Hybridized Electric Auto-Rickshaw under Dr. Ojapah Mohammed

Digital Skills

Programming Languages	Python (Pyspark included), SQL (Postgres), SparQL
Scripting and Version Control	Visual Studio, Jupyter, Kaggle, Google Colab, Atom, Git
Database Creation	AWS RDS, PGAdmin, IBM DB2, ERDPlus, Azure Database, GCP SQL
Cloud and Container Services	AWS, Microsoft Azure, Google Cloud Platform, IBM Cloud, Heroku, Digital Ocean, GitHub Action, Docker
Visualization	Matplotlib/Seaborn, Power BI, Tableau, Excel
OS	Windows, Linux, Ubuntu
Soft Skills	People Management, Time Management, Teamwork, Problem-solving, Consistency, Documentation

Work Experience

Knowledge Beyond Borders Africa

Abuja, Nigeria

Data Scientist/Analyst

Sep 2019 - Aug 2021

- I worked with my manager to apply descriptive analysis, factor analysis and hypothesis testing on research data, in order to proffer solution to the downtime experienced by MTN communications system, amongst numerous other clients.

Upwork Freelance

Global Remote

Applied Data Scientist

Feb 2022 - Present

- **Certificate:** [Link to Profile](#)
- **Specialties:** Advanced ETL Pipeline Design and Cloud Deployment, Natural Language Processing, LLM (GPT, Llama2) Integration, Machine Learning Model Development and Cloud Deployment, AI End-to-End Solution Implementation

Notable Projects

Alfred (AI Powered Assistant) - Ongoing

L'Aquila, Italy

Personal Project

2023

- This project will involved the use of combinatorial optimization, network analysis and the integration of LLMs (GPTs) to provide assistance to users

GPT-Powered Automated Data Analyst

L'Aquila, Italy

Personal Project

2023

- This Project is about automating (using GPT) the processing of a source .csv file. It compares each column in the source file to the columns in the template .csv file and uses the most similar columns in the template file to generate python codes and uses these codes to automatically transform the columns in the source file. It also discards duplicate columns and irrelevant columns in the source file.
- GitHub Repo:** Repository Link, Loom Video
- Tools Used:** Streamlit, OpenAI GPT-3 API, Prompt Engineering

Citizens' Voice Platform

L'Aquila, Italy

Personal Project

Aug 2022 - Mar 2023

- This project is the proof of concept (POC) of a bigger idea, which is based on raising awareness amongst the people on their rights and powers as "the people", helping them understand the government structure, get to know their leaders, and make their voices a lot more audible. It focuses on visualization of citizens' sentiment and trending discussions (using PowerBI dashboard) which is automatically and periodically predicted (using NLP, Tensorflow, Pytorch SimpleRNN, LSTM, BERT etc.) and extracted (using Gensim LDA) respectively, for technical and non-technical users. This data is automatically and periodically extracted (using Twitter API and GitHub Actions) from citizens' social platforms and stored using AWS RDS. These results and data are also made available to technical users using a REST API and Python Third Party Library (TPL). Deployment was done on Heroku, Pypi, AWS, Digital Ocean and Docker.
- Report Link:** Link to Conclusion Report Posted on LinkedIn, Link to Report on GitHub
- Dashboard:** Link to Published PowerBI Dashboard
- Phase 1 Repo:** Link to Phase 1 GitHub Repository, Link to Phase 1 LinkedIn Article
- Phase 2 Repo:** Link to Phase 2 GitHub Repository, Link to Phase 2 LinkedIn Article
- Tools Used:** AWS RDS, PGAdmin4, TwitterAPI, FastAPI, VSCode, Heroku, Pypi, BERT, Gensim, Tensorflow, Pytorch, Scikit-learn, Docker, Digital Ocean (Kubernetes), GitHub Action, Git, Ubuntu 22.04, GPU Tesla T4, PowerBI

PDF Reader

L'Aquila, Italy

Personal Project

2023

- This is a brief project which extracts tables from PDF files, embeds the contents of the table using OpenAI LLM and upserts the contents into a Vector Store.
- GitHub Repo:** Repository Link, Loom Video
- Tools Used:** Streamlit, Camelot, Weaviate and OpenAI

UK Unemployment Rate Time Series Project

L'Aquila, Italy

Universita' degli Studi dell'Aquila

Dec 2022

- This project focuses on gathering time series data for the UK Unemployment Rate, transforming the data, building a suitable ARIMA model for the data, forecasting future values using the built model, and comparing this forecast with the real world outcome.
- GitHub Repo:** Repository Link
- Tools Used:** Gretl, Excel

Selection Methods Python TPL (Contribution to Python Open Source)

L'Aquila, Italy

Personal Project

Aug 2022

- Third Party Library (TPL) for selecting the optimal features using the various selection methods including stepwise algorithm amongst others.
- Project Links:** Library Link on Pypi, GitHub Repository Link
- Tools Used:** VS Code, Git, Pypi

My Portfolio

-

Content Description

2023

- My portfolio contain a wide variety of projects ranging from simple projects to advance end-to-end solutions, involving the use of various machine learning methods, as well as project done in the Forex/Trading sector and projects carried out in my previous engineering specialty.
- Portfolio link:** GitHub Repository Link

Publications

PAPER(S)

Michael Dubem Igbomezie, Phuong T. Nguyen, Juri Di Rocco, Davide Di Ruscio, CO3D: Fusion of Word Embedding and LSTM for

In Review Detecting Code Comments Coherence, demo at the 38th IEEE/ACM International Conference on Automated Software Engineering
New Ideas and Emerging Results track (ASE 2023 NIER) (ASE 2023 NIER)

Awards

2021 **Double Degree Scholarship**, InterMaths Consortium

Italy

Languages

English Bilingual proficiency

Italian Basic proficiency

Authorization to process my personal data _____

Pursuant to Legislative Decree 196/03 and the GDPR (EU Regulation 2016/679), I authorize any interested company or her representative to process my personal data for the purpose of this application.

Note: Many other Data Science projects (involving customer segmentation, statistical third party library development and deployment etc.) can be found on my GitHub profile. I am also open to relocation and traveling.

References available upon request.