Michael Dubem Igbomezie

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

+39 350 103 0134 | michaeligbomezie@gmail.com | michaeligbomezie.wixsite.com/portfolio | 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.

Work Experience

Upwork Freelance   *Global Remote*

Applied Data Scientist (**Top Rated Profile**, **100% Job Success**)   *Feb 2022 – Present*

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

- Developed a script that conducts various types of analyses on video data (including text, audio, and image analysis) using various publicly available APIs. This project involved audio transcription, text analysis using LLMs, image analysis using LVMs, parallel computation etc. I also developed a version of the script to carry out analysis on large volumes of videos (in millions) in reasonably short periods of time.

- Developed an intelligent web scraper, which extracts information from websites independent of the website structure.

- Developed a predictive system that automatically extracts sports related data from multiple websites (static and dynamic), processes and transforms this data and stores it in a hosted database. This transformed data is then extracted and fed into two different types of predictive models, to predict the outcome of sport matches. This project also involved historical analysis of sports data to determine the efficiency of the developed system

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.

Education

Universita' degli Studi dell'Aquila   *L'Aquila, Italy*

MSc in Applied Data Science   *Sept 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
Distributed Computation	I/O and CPU Bound Distributed Computation, Apache Spark
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

Notable Projects

Predicting Terrorist Attacks in Nigeria - Ongoing
L'Aquila, Italy

Humanitarian Project
2024 - Present

- This project is aimed at extracting publicly available historical data from the internet from various news websites. Then LLMs will be used to conduct feature engineering and extraction to build a dataset that can be analyzed. EDA will be conducted to discover any patterns. Time series models and other non-linear models will be built to predict the time, frequency and location of terror attacks.

Alfred (AI-Powered Assistant) - Ongoing
L'Aquila, Italy

Personal Project
2023 - Present

- This project is a virtual assistant built to interpret building plans and provide directions and assistance to users using combinatorial optimisation, network analysis and the integration of LLMs.
- **GitHub Repo:** Repository Link
- **Tools Used:** OpenAI GPT-3, LangChain, Pinecone Vector Store, Telebot, Prompt Engineering

Citizens' Voice Platform
L'Aquila, Italy

Humanitarian Project
Aug 2022 - Mar 2023

- A proof of concept (POC) focused 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 visualizes the citizens' sentiment and trending discussions which is automatically and periodically predicted (using NLP, Tensorflow, Pytorch SimpleRNN, LSTM, BERT, Gensim LDA, PowerBI dashboard etc.) respectively for technical and non-technical users. Using Twitter API and GitHub Actions, data is extracted from citizens' social platforms and stored in an AWS RDS in a scheduled manner. These results and data are also made available to technical users using a REST API and Python Third Party Library. 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

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

My Portfolio
-

Content Description
2023

- My portfolio contains a wide variety of projects ranging from simple projects to advanced 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)

In Review Michael Dubem Igbomezie, Phuong T. Nguyen, Juri Di Rocco, Davide Di Ruscio, *CO3D: Fusion of Word Embedding and LSTM for Detecting Code Comments Coherence*, demo at the 28th International Conference on Evaluation and Assessment in Software Engineering (EASE 2024)

Awards

2021   **Double Degree Scholarship**, InterMaths Consortium
Italy

Languages

English   Bilingual 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.

References available upon request.