Bright Uzosike

Data Scientist · ML Engineer · Manchester, UK · hello@brightuzosike.dev · brightuzosike.dev · linkedin.com/in/bright-uzosike · github.com/michizler

SUMMARY

Data scientist who builds, ships and maintains explainable ML systems that add business value. I evaluate before I trust, so gold sets, cross-validation and a written list of limitations ship with every model. I hold an MSc in Data Science with Distinction from the University of Salford and have six years of experience across analytics, BI and machine learning.

EXPERIENCE

Research Data Scientist · Rebel Hive — May 2026 to present

- Sole data scientist on a live client project, taking an ambiguous brief to an NLP diagnostic tool now in commercial use by the client.
- Engineered a 76-feature matrix from 38 regulatory and corporate documents. Classifiers predicting Ofsted outcomes from language alone reached macro-F1 of 0.82 to 0.91 under 5-fold stratified CV on 24 inspection reports.
- Validated with pre-specified hypotheses, non-parametric tests, robustness checks and an inter-rater reliability harness, and reported counter-theoretical findings rather than overclaiming.
- Shipped a single-command pipeline with a pinned environment and fixed seeds, served through FastAPI and Streamlit, with a replication guide and a forward data contract.

AI Data Specialist (RLHF & SFT) · Mercor — Aug 2026 to present

- Score and rank 2 to 8 candidate model outputs per prompt against multi-axis rubrics, writing auditable rationales that feed reward-model training.
- Flag reward-hacking failures where polish hides substance, such as fabricated tables, unsupported claims and arithmetic errors.
- Build Azure Data Factory pipelines for streaming recommendation systems, and Informatica DEI mappings, to per-task SLAs.

AI Training Contributor (SFT & RLHF) · Outlier — Sep 2025 to present

- Passed the Generalist SFT and RLHF screening. Evaluate model outputs for accuracy, instruction-following and safety, writing structured feedback used in retraining.

Data Science Consultant (internship) · Amdari — Jan 2025 to Jun 2026

- Delivered end-to-end ML engagements for several clients, covering scoping, EDA, feature engineering, modelling, evaluation (precision, recall, F1, AUC-ROC) and stakeholder reporting.
- Built ETL/ELT pipelines in Azure Data Factory, and agentic AI and RAG automations on Microsoft Power Platform.
- Turned model outputs into dashboards and presentations for non-technical stakeholders. Retained beyond the initial term to deliver for two further clients.

GIS Analyst · Nationhood Podcast — Mar 2025 to Apr 2025

- Georeferenced and digitised map files (GeoTIFF, GeoPackage) in QGIS and prepared them for web embedding.

EDUCATION

MSc Data Science, Distinction

University of Salford · 2025 to 2026

Dissertation on an NLP diagnostic of organisational language in further education, delivered as a live client project.

BEng Electronic Engineering

University of Nigeria, Nsukka · 2018 to 2024

Second Class Upper, GPA 4.45/5.00.

SKILLS

Languages: Python, SQL (T-SQL, PostgreSQL, BigQuery, Spark SQL), R, JavaScript, C, Bash

Modelling and evaluation: scikit-learn, LightGBM, XGBoost, statsmodels, SHAP, cross-validation, gold-set evaluation, time series (ARIMA, GARCH)

NLP and LLMs: spaCy, Hugging Face Transformers, Claude API, Pydantic schemas, RAG, RLHF and SFT

Data engineering: Azure Data Factory, Informatica DEI, Databricks (Spark, MLlib), Microsoft Fabric, ETL/ELT

BI: Power BI (DAX), Tableau, Excel, Streamlit

Shipping: Docker, Docker Compose, Git and GitHub, CI/CD, Linux, Render, MLflow

CERTIFICATES

Google Data Analytics Professional Certificate (2022)

Share Data Through the Art of Visualization, Google (2022)

ALX Data Analytics Certificate (2025)

LANGUAGES

English (native or bilingual), Spanish (elementary)

HR Data Analyst to Branch Director · AIESEC Nigeria & UK

Apr 2022 to Jul 2025

- Rose from Talent Manager and HR Data Analyst to Branch Director, leading distributed teams of 50+ members, then led the HR and organisational development team at AIESEC UK.
- As Branch Director, led AIESEC in Enugu to the top spot in operations out of 13 branches in Nigeria in July 2024.
- Built KPI dashboards for all 13 national branches and raised data consistency by 10% by linking reporting to cleaned source data.
- As HR Team Lead, raised data integrity to 93% through validation and review of the team's datasets.

Business Intelligence Analyst · TRACKER.NG Jan 2023 to Aug 2023

- Automated Tableau dashboards joining sales, customer service and operations data, cutting a decision-making bottleneck from 35% to 15%.
- Added validation logic that raised data integrity by 30% and halved ad hoc correction requests.
- Liaised between business development, operations and IT, cutting project delivery timelines by 20%.

Software Engineering Intern · ALX Africa May 2022 to Jul 2023

- Completed a 12-month software engineering programme, building full-stack Python and JavaScript applications on Linux.

Data Analyst Intern · Trainiton Analytics Oct 2020 to Apr 2022

- Automated repetitive data processing in Python, and co-led analyst training with 85% retention over the final six months.

SELECTED PROJECTS

AlloyTower AVM

8.0% MAPE with an assessment, 44.9% without (5-fold CV). Every estimate is explained with SHAP and served by a five-endpoint FastAPI service in Docker that redeploys on every push.

VoyageRail fare desk

Fare model that quotes from what a booking desk knows at quote time and explains every price with SHAP. Gradient boosting reaches R^2 0.90 against a 0.71 linear baseline.

LLM Feedback Analyzer

Schema-validated Claude output with retries, backoff and a safe fallback. 13 of 15 correct on a hand-labelled gold set for sentiment and churn risk.

US Bancorp volatility

Holt-Winters and ARIMA failed Ljung-Box. ARIMA(1,0,1) with sGARCH(1,1) passed ($p = 0.2369$) and gave a 365-day price and volatility forecast.