PRINCE OKUNADE

Data Scientist | Machine Learning Engineer

London, UK | 07350153650 | princeokunade1@gmail.com | LinkedIn | GitHub | No sponsorship required

PROFESSIONAL SUMMARY

Data scientist and aspiring ML/software engineer with a statistics foundation, working across the full lifecycle from data pipelines to deployed, instrumented models. I build and evaluate LLM and machine learning systems rigorously, with a strong bias toward testing, measurement, and quality: one controlled change at a time, documented results, and code deployed and monitored in production-style environments. Strong command of Python, SQL, Git, Docker, and cloud tooling, with a solid grounding in statistics including a co-authored published Bayesian study, and a track record of shipping projects end to end.

TECHNICAL SKILLS

Languages & Tools: Python, R, SQL, Git, Jupyter

ML & AI: Scikit-learn, XGBoost, LightGBM, TensorFlow, SHAP, Hugging Face Transformers, LLMs (Claude API), retrieval-augmented generation (RAG), NLP

Data Science Methods: Feature Engineering, Model Evaluation, Cross-validation, Experimental Design, Statistical Modelling, Time-Series Analysis, A/B Testing

Cloud & MLOps: GCP (Cloud Run, BigQuery, Artifact Registry, Cloud Build), Azure ML Studio, Hugging Face Spaces, Streamlit, Docker, GitHub Actions, CI/CD, FastAPI

Data & Visualisation: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Power BI, Excel

PROJECTS

Retrieval-Augmented QA over UK Drug-Safety Updates — Grounded Drug-Safety RAG [GitHub] [Live demo] [Article]

- Built and deployed a live RAG system over UK drug-safety documents that enforces faithfulness as a safety property: 100% faithfulness and 0 false answers on 16 adversarial questions.
- Raised retrieval from a 70% baseline to 81% recall@1 via one-variable experiments (hybrid dense + BM25, reciprocal rank fusion), each gated on a self-authored 80-question benchmark scored by a validated LLM-as-judge.
- Cut a reranker that added 15x latency for no quality gain and documented the negative results in a published write-up. Measured, not assumed.

Customer Churn Prediction — Containerised ML Deployment on GCP [GitHub]

- Turned a churn model into a prioritised retention call list with 2.5x lift: the top 10% of customers reaches roughly 2 in 3 real churners, and the top 30% captures 63% of all churners.
- Served per-prediction SHAP explanations through the API, logged predictions to BigQuery, and used SQL to monitor risk-level distribution and churn drift against the training rate.

Stock Price Direction Predictor — Live API-Driven ML Pipeline [GitHub]

- Built an end-to-end pipeline predicting next-day price direction from live Yahoo Finance data, engineering 20+ technical indicators (RSI, MACD, Bollinger Bands, ATR, OBV) across Logistic Regression, Random Forest and XGBoost.
- Tuned XGBoost with GridSearchCV under TimeSeriesSplit cross-validation to prevent look-ahead bias, with a Streamlit dashboard serving live predictions, SHAP explanations and a strategy backtest.

EDUCATION

MSc Data Science — York St John University, London Sep 2025 – Sep 2027

- Two-year programme with a 12-month industry placement scheduled June/September 2026 – 2027.
- Coursework: AI and ML, Cloud Computing (Azure, GCP), Data Systems and Security, Blockchain Development.

BSc Statistics — University of Ibadan, Nigeria Jun 2019 – Apr 2024

- Thesis: Bayesian Analysis of COVID-19 Recovery and Death Cases in Nigeria (Weibull-Gamma and Lognormal-Inverse Gamma survival models across 37 states).
- Publication: Akanbi, O. B., and Okunade, P. O. (2026). Bayesian Analysis of COVID-19 Recovery and Death Cases in Nigeria. *Journal of Scientific Research and Reports*, 32(6), 29–44. https://doi.org/10.9734/jsrr/2026/v32i64223

PROFESSIONAL EXPERIENCE

Data Science Intern, Biometrics Unit — IITA, Nigeria Jun 2023 – Sep 2023

- Cleaned, validated and analysed agricultural field-trial data from 8+ experimental locations in R, preparing 15,000+ observations for crop-yield and biometrics analyses.
- Ran exploratory analysis and fitted linear and logistic regression models as part of statistical modelling and biometrics training.

Mathematics Tutor — Get Further, London, UK Jan 2026 – Present

- Delivering targeted GCSE maths tuition to 20+ further-education students within a nationally recognised resit programme, supporting problem-solving and exam readiness.