

ONUCHUKWU JOSEPH CHIMEZIE

AI and Data Science Researcher | Machine Learning Practitioner

prolifiqmexie@gmail.com | linkedin.com/in/onuchukwu-joseph-589912148 | github.com/prolifiq1 |
chimezie-ai-portfolio.netlify.app

PROFILE

Data scientist and AI practitioner with demonstrated experience in deep learning, classical machine learning, statistical analysis, and data-driven decision-making. Strong focus on rigorous model evaluation, reproducible methodology, and translating quantitative findings into actionable insights. Seeking doctoral research opportunities at the intersection of applied AI, model reliability, and real-world intelligent systems.

EDUCATION

MSc Data Science

University of Hull, United Kingdom | 2024 - 2025

- Core modules: Understanding AI, Fundamentals of Data Science, Programming, Data Visualisation, Advanced Analytics
- Portfolio work spanning convolutional neural networks, supervised and unsupervised learning, demographic data analysis, and applied statistical modelling
- Developed complete ML pipelines with systematic experimentation, hyperparameter tuning, and comparative evaluation frameworks

BSc Computer Science

Undergraduate Degree

- Foundation in algorithms, data structures, software engineering, and computational thinking

SELECTED TECHNICAL PROJECTS

Vehicle Damage Classification Using Convolutional Neural Networks

Deep Learning | Computer Vision | TensorFlow/Keras

- Designed and trained a CNN to classify six categories of vehicle damage from photographs for insurance claim verification
- Built a baseline architecture, then systematically improved generalisation through dropout, L2 regularisation, data augmentation, and hyperparameter tuning
- Addressed class imbalance using computed class weights and stratified splitting
- Conducted overfitting analysis comparing training and validation dynamics across experiments

Streaming Service Customer Behaviour Prediction

Machine Learning | Comparative Study | scikit-learn / TensorFlow

- Built end-to-end ML pipeline on 5,000 customer records: regression (linear, polynomial, Random Forest, ANN), classification (Logistic Regression, Random Forest, SVM), and clustering (k-Means, Agglomerative)
- Established shared preprocessing pipelines and fixed train-test indices to ensure fair cross-method comparison
- Evaluated models using RMSE, R-squared, ROC-AUC, silhouette scores, and per-class metrics
- Identified Random Forest with mixed features as the strongest regressor; produced actionable customer segmentation profiles

Census Demographic Analysis and Policy Recommendations

Data Science | Statistical Analysis | Pandas / Matplotlib

- Cleaned 8,175 census records using household-aware imputation methodology, legal precedent, and demographic context

- Estimated crude birth, death, and migration rates from cross-sectional age structure data
- Performed statistical testing (t-tests, chi-squared) comparing town demographics against England and Wales national benchmarks
- Produced evidence-based policy recommendations for land use and public investment, with all alternative options evaluated and rejected on statistical grounds

TECHNICAL SKILLS

Machine Learning	Regression, Classification, Clustering, Neural Networks, Model Selection, Hyperparameter Tuning, Feature
Deep Learning	CNNs, Transfer Learning, Image Classification, Data Augmentation, Regularisation (Dropout, L2, Early Sto
Data Science	EDA, Statistical Testing, Data Cleaning, Demographic Analysis, Data Visualisation, Reproducible Analysis
Programming	Python (NumPy, Pandas, scikit-learn, TensorFlow/Keras, Matplotlib, Seaborn), SQL, Jupyter, Git

RESEARCH INTERESTS

- Applied computer vision: generalisation beyond benchmark datasets in practical domains
- Model reliability and evaluation: uncertainty quantification, trustworthy AI, the gap between validation metrics and deployment performance
- Data-driven decision support: demographic modelling, resource allocation, responsible translation of analysis into policy
- Interpretable and responsible AI: fairness, transparency, and auditability in automated decision-making systems

KEY STRENGTHS

- Systematic experimental methodology with controlled comparisons across models and methods
- Domain-informed data preparation that goes beyond mechanical preprocessing
- Clear technical communication with thoroughly documented reasoning
- Ability to connect quantitative findings to real-world decisions and actionable recommendations
- Honest evaluation that prioritises reliability and reproducibility over headline metrics