

ABDULLAH AL MAMUN

Junior Data Scientist | Python, Machine Learning, Statistical Modelling & Data Visualisation

Bangor, Gwynedd | +44 7481 005465 | mamunabdulla40@gmail.com | linkedin.com/in/aamamun234 | github.com/aamamun1 |
www.aamamun.com

PROFESSIONAL PROFILE

Data Scientist with an MSc (Merit) in Computing for Data Science and hands-on experience building end-to-end machine learning and analytics workflows in Python. Experienced in data cleaning, feature engineering, regression modelling, statistical validation, model comparison, error analysis, and insight-led visualisation. Applied project work includes manufacturing analytics, climate trend analysis, and audio similarity detection. Known for reproducible workflows, clear documentation, limitation-aware reporting, and translating complex results into practical, decision-ready summaries.

CORE DATA SCIENCE SKILLS

Programming & Tools:	Python, Jupyter Notebook, VS Code, Git, LaTeX, C (foundational), database fundamentals
Data Science & ML:	Data cleaning, EDA, feature engineering, supervised learning, regression, model selection, cross-validation
Evaluation & Statistics:	Error analysis, statistical validation, correlation analysis, trend analysis, assumptions tracking, metric reporting
Libraries:	Pandas, NumPy, scikit-learn, Matplotlib, Seaborn, SHAP, XGBoost, Librosa
Visualisation & Reporting:	Clear charts, publication-quality figures, comparative plots, heatmaps, interpretable summaries, limitations-first reporting
Working Style:	Reproducible pipelines, notebook hygiene, clear README notes, concise communication, incremental model improvement

EDUCATION

MSc Computing for Data Science (Merit) - Bangor University, United Kingdom 12/2023	09/2022 -
• Key areas: Machine Learning, Predictive Analytics, Statistical Modelling, Data Analysis and Visualisation, Research Methods, and Python for Data Science.	
BSc Computer Science & Engineering - Dhaka International University, Bangladesh 12/2019	04/2016 -
• Core training: Programming, Algorithms, Software Development, Databases, Embedded Systems, and final-year robotics project using Arduino Mega and C.	

SELECTED DATA SCIENCE PROJECTS

MSc Dissertation - Predictive Modelling for 3D-Printed Metal Components • Developed a full predictive modelling workflow to estimate yield strength from additive manufacturing process parameters. • Prepared data through missing-value handling, consistent splitting, standardised inputs, and repeatable experiment structure. • Built and compared regression approaches using objective evaluation, model selection, validation, and error analysis. • Investigated process drivers such as laser power and scan speed to support process-property understanding and practical interpretation. • Produced clear visual summaries and dissertation-ready reporting covering findings, comparisons, limitations, and reproducibility.	08/2023
Climate Data Analysis with Machine Learning • Analysed long-term climate datasets to identify temperature and rainfall trends over time. • Applied regression modelling, statistical validation, yearly/seasonal aggregation, and evidence-led interpretation. • Created repeatable analysis pipelines and interpretable time-series/comparative plots that could be regenerated from raw data.	
Voice Similarity Detection System (MFCC + Cosine Similarity) • Developed a Python-based system to measure and visualise speaker similarity for comparison tasks. • Extracted MFCC features, calculated cosine similarity scores, and generated interpretable visual outputs for multiple sample pairs.	
Model Robustness & Evaluation Practice (Reusable Template) • Developed a reusable evaluation template covering data split discipline, cross-validation, metric reporting, sanity checks, and assumptions tracking.	

- Included reliability checks such as outlier sensitivity, feature leakage awareness, and stability across folds.
- Emphasised repeatable results and clear reporting for portfolio-grade data science deliverables.

Data Visualisation & Reporting Pack

- Created structured visual outputs for projects, including trend plots, comparison plots, error plots, and heatmaps.
- Standardised figure layouts and captions to support professional reporting, academic-style presentation, and clear stakeholder interpretation.
- Focused on interpretability by presenting results with context, limitations, and actionable takeaways.

Humanoid Robot Control Project (DOF-17 Robot)

12/2019

- Designed and programmed coordinated motion control routines and basic interaction logic using Arduino Mega and C.
- Integrated sensors and actuators, applied calibration and iterative testing, and improved repeatability in movement sequences.
- Strengthened understanding of real-world constraints including timing, stability, power, calibration, and repeatable execution.

PUBLICATION

Predicting Yield Strength of 3D-Printed Metal Components Using Machine Learning and Process Parameters

02/2026

- First author conference paper published in Springer Lecture Notes in Networks and Systems, Proceedings of the 3rd International Conference on Big Data, IoT and Machine Learning, pp. 490-498. DOI: 10.1007/978-3-032-15346-3_34.

ROLE ALIGNMENT: JUNIOR DATA SCIENTIST

- Comfortable working with Python-based analysis workflows from raw data preparation to model evaluation and clear reporting.
- Strong foundation in basic modelling techniques, including regression, supervised learning, baseline comparison, cross-validation, and metric-led evaluation.
- Able to explain findings clearly using visual outputs, concise summaries, and limitation-aware communication for non-technical audiences.
- Interested in applying data science to operational, customer, and commercial questions through practical analysis, experimentation support, and decision-ready outputs.

EVIDENCE FOR THE JOB REQUIREMENTS

- SQL/Python/data handling: CV evidence includes database fundamentals, Python workflows, Pandas, NumPy, structured data preparation, and repeatable analysis pipelines.
- Data cleaning and analysis: project evidence includes missing-value handling, consistent splitting, standardised inputs, EDA, correlation analysis, and trend analysis.
- Foundational models: project evidence includes regression modelling, supervised learning, baseline comparison, model selection, cross-validation, and metric-led evaluation.
- Visualisation and reporting: project evidence includes Matplotlib, Seaborn, trend plots, comparison plots, error plots, heatmaps, figure captions, and clear written summaries.
- Business communication: CV evidence shows limitation-aware reporting, concise summaries, practical interpretation, and explaining results for non-technical readers.

ADDITIONAL TECHNICAL STRENGTHS

- Reproducibility: consistent project structure, notebook hygiene, README notes, documented assumptions, and repeatable runs.
- Communication: concise summaries, visual explanation of results, practical recommendations, and transparent discussion of limitations.
- Engineering discipline: baseline-first modelling, incremental improvements, version control with Git, and documented metric tracking.