Callum Leach

Residence/domicile: Devon

E-mail: callumleach31@gmail.com ✻ *Telephone number:* +44 07503 292191

Professional Profile

Mathematics graduate driven by a passion for applying mathematical knowledge to real-world challenges. Recognised the power of data at Eaton Aerospace and later expanded expertise through various collaborative projects. Now eager to leverage analytical, technical, and data-driven skills built within shell to explore how data powers the future.

Career Summary

Virgin Atlantic *April 2025 -*
Data Scientist *London, UK*

- Developed AMIE (Aircraft Maintenance Insights Engine) in Databricks and Streamlit to cluster recurring aircraft defects, enabling proactive maintenance and cost savings.
- Built machine learning pipelines using text embeddings, KMeans clustering and LLM labelling to rapidly detect high-cost recurring defects from unstructured engineering logs.
- Evaluated and monitored model performance using silhouette and other clustering metrics, tracking stability and drift over time to ensure reliable insights.
- Delivered dashboards tracking defect frequency and trends, supporting engineering teams in reducing downtime and improving fleet reliability.
- Developed a hurdle regression model in Databricks with MLflow to predict airport arrival delays, combining logistic regression (for zero vs non-zero delay) and ridge regression (for magnitude), handling the zero-inflated nature of the target.
- Enabled data-driven fuel planning, reducing waste from manual buffer estimates and delivering potential savings of more than £0.6M annually.

Shell *December 2022 - December 2024*
Data Scientist *London, UK*

- Collaborated with dynamic team to deploy Azure Databricks pipelines for offshore wind-farm predictive maintenance, implementing quantile regression and statistical measures to pioneer an AI solution predicting turbine failures up to 6 months in advance, reducing downtime and elevating operational efficiency.
- Designed a Plotly Dash dashboard for targeted and intelligent wind turbine failure detection, empowering operations teams to diagnose issues quickly and efficiently.
- Developed & deployed a scalable package in PySpark for degradation analysis of green hydrogen electrolysers across multiple assets. Used statistical methods such as Kernel density estimation and distance metrics to determine degradation under grid and photovoltaic operation.
- Conducted engaging "Data in Action" presentation for 100+ participants at a data boot camp, emphasising the pivotal role of data analytics in driving business success.
- Alongside university researchers I actively co-authored a research paper on extreme value analysis of climate models. Employing cutting-edge statistical techniques like Bayesian inference on peaks-over-threshold and block maxima. This provided actionable insights that guided strategic decision-making within the de-carbonisation team at Shell.
- In collaboration with the Inner-source team, I developed a GitHub bot to assist developers in reaching Inner-source standards. This increased code maturity and achieved code compliance of 80%.

Kubrick Group *August 2022 -*
Data Engineer *London, UK*

- Orchestrated ETL and CI/CD pipelines in various cloud based systems during my training.
- Developed alongside Mental Health Innovations a containerised document redaction API using Docker.

Eaton Aerospace *June 2017 - September 2019*
Manufacturing Engineering Technician *South Molton, UK*

- Programmed robots in Java to assemble actuators and utilised cameras and neural networks to perform simple non-conformance detection.
- Used statistical method such as Weibull distributions and survival functions to help the maintenance department take a more proactive approach to minimising machine downtime.
- Had the opportunity of working with multiple manufacturing & engineering departments. Engaged in a range of data collection, analysis, artificial intelligence, robotics and industry 4.0 initiatives in a manufacturing environment.

Education

University of Bath *September 2019 - June 2022*
BSC - Mathematics *Bath, UK*

- Completed masters modules in Bayesian statistics & mathematical modelling.
- Completed Echo State Network project to model chaotic dynamical systems.

Exeter Tutorial College *September 2017 - July 2019*
A-Levels *Exeter, UK*
Mathematics, Further Mathematics, Physics, Biology
Edexcel Advanced Extension Award (AEA) in Mathematics