

ZHIFENG LI

Applied Machine Learning Engineer | Research Engineer

zl5825@ic.ac.uk | github.com/ada-zl5825

PROFILE

Applied ML and research-software engineer completing Imperial's MSc in Applied Computational Science and Engineering, with a First Class BSc in Computer Science. Builds reproducible systems spanning local RAG/SLM evaluation, rare-event ML, sensor fusion, scientific computing and parallel C++/MPI.

EDUCATION

Imperial College London

MSc Applied Computational Science and Engineering | Sep 2025 - Sep 2026 (expected)

- Scientific computing, numerical methods, optimisation, deep learning, advanced programming and parallel programming.

University of Nottingham

BSc Computer Science, First Class | Sep 2021 - Jun 2025

- 2+2 programme across Ningbo, China and the UK; coursework included operating systems, computer vision, algorithms, databases and distributed systems.

TECHNICAL SKILLS

ML and evaluation: PyTorch, scikit-learn, reproducible train/validation pipelines, weighted and focal losses, precision/recall/F1 diagnostics

RAG and LLM evaluation: Foundry Local, local SLMs, chunking, embeddings, vector/hybrid retrieval, citation correctness, groundedness, abstention and regression tests

Scientific and software engineering: Python, C++17/20, ODE integration, finite differences, optimisation, numerical linear algebra, Linux, MPI, OpenMP, CMake, Git, pytest, Sphinx, Streamlit

EXPERIENCE

Microsoft Project via Imperial College London

Applied ML Research Engineer Intern (MSc Independent Research Project) | 2026 - Present

- Building a local-first RAG pipeline over public AML/CDD guidance, synthetic policies and synthetic onboarding cases, with local inference after initial setup.
- Designing labelled JSONL benchmarks for retrieval, citation validity, groundedness, abstention, task success and cold/warm latency; comparing RAG with keyword and no-retrieval baselines.
- Evaluating chunk size, overlap, top-k, metadata filters and embedding choices without presenting target thresholds as completed results.

Hefei TYDIC Information Technology Co., Ltd.

Big Data Development Engineer Intern | Jun 2023 - Aug 2023

- Built operations and maintenance scripts that reduced repetitive monitoring work and supported internal platform reliability.
- Supported deployment, configuration and incident troubleshooting for internal application platforms; documented procedures for engineering handover.

SELECTED PROJECTS

Predicting the Unpredictable - Storm Forecasting and Lightning Prediction

Nine-person MSc team coursework | PyTorch, rare-event ML, reproducible evaluation

- Contributed to a shared package spanning four tasks over 800 storms: 12-frame VIL forecasting, satellite-to-VIL reconstruction, eight-class event classification and lightning prediction.
- The team artefact standardised data, preprocessing, metrics, plotting and submission I/O, with pytest and Sphinx support and imbalance-aware loss/metric workflows.

ArUco + MediaPipe Hand Tracking Fusion

Computer vision prototype | OpenCV, MediaPipe, Kalman filtering

- Built a real-time prototype combining calibrated ArUco pose estimates, MediaPipe landmarks and Kalman-based fusion, with camera, video and headless capture modes.
- Structured acquisition, detection, fusion and utilities behind one CLI; added CSV/video logging, selectable comparison modes and optional frame-aligned ground truth.

Scientific Computing and Optimisation

DeepImpact | Genetic Algorithm Circuit Optimizer

- DeepImpact: implemented atmospheric-entry ODE simulation and post-impact risk mapping linked to UK postcode and population exposure data, with reusable Python/Streamlit modules and edge-case tests.
- Circuit Optimizer (team coursework): contributed to a C++20 system combining genetic search, circuit simulation, structural validity checks and reproducible seeded benchmarks.

ADDITIONAL WORK

Medical AI coursework: PyTorch MRI/ECG modelling with preprocessing, training, validation and failure-mode review.

Parallel Navier-Stokes: C++17/MPI solver with bit-exact serial/cross-rank regression checks and strong-scaling experiments to 128 ranks.