Matthew Archer

Senior ML/Research Software Engineer

Computational Physicist and Senior Research Software Engineer with 7+ years' experience bridging numerical simulation and scalable Deep Learning. Specialist in coupling ML surrogates with Fortran/C++ scientific codes (e.g., FTorch) and optimising distributed training on heterogeneous HPC architectures. PhD in Physics with deep expertise in FVM and FEM solvers.

Experience

Senior ML/Research Software Engineer | Institute of Computing for Climate Science (ICCS) |University of Cambridge, UK | Jan 2024 - Present

I specialise in applying machine learning and high-performance computing techniques to solve complex modelling problems in climate science.

- Led a team of 3 engineers to design and deliver a modular forecasting and physical evaluation framework of AI-based emulators for ocean spin-up workflows.
- Supported the design and integration of ANN and UNet neural network emulators (local and non-local) for gravity wave parameterisation in atmospheric models.
- Supervised and mentored junior engineers, establishing best practices for ML engineering (evaluation, testing, reproducible pipelines). Taught fundamentals of Python, PyTorch and ML to summer-school delegates.

ML/Research Software Engineer | UIS, University of Cambridge, UK | Dec 2017 - Dec 2023

Contributed to large-scale machine learning and scientific computing projects, with a focus on scalable training and HPC.

- Built and evaluated large-scale ML pipelines across HPC systems, including web-scale NLP workflows using Common Crawl data (multi-terabyte scale). Explored and compared GraphSAGE, Doc2Vec, and TF-IDF embeddings for downstream document and graph classification tasks.
- Worked on the performance optimisation and scaling of deep learning models (e.g., AlphaFold2, vision, and LLM inference) across heterogeneous hardware platforms and interconnects.
- Developed a hybrid FEM-BEM solver for wave propagation using FEniCS and BEMpp. Designed a parallel architecture using MPI communicator splitting (32:1) to balance load between coupled solvers.
- Set up CI pipelines and reproducible ML/HPC benchmarking infrastructure FEniCS benchmarking using GitLab CI and reframe, to support model and solver development on pre-exascale supercomputers.

Technical focus

- Large-scale ML & Distributed Training and inference across heterogeneous hardware. Benchmarked NCCL AllReduce performance across different interconnects (NVLink vs. PCIe) and tested data staging, mixed precision etc.
- Reproducible ML Engineering: CI/CD for ML/HPC workflows (GitLab CI, ReFrame), benchmarking on pre-exascale systems, robust evaluation pipelines.

Publications

- **M. Archer** (contributor). Tree crown detection using Mask R-CNN. Remote Sensing in Ecology and Conservation, 2023. [paper, code]
- **M. Archer** (contributor). Pushing the limits of exoplanet discovery via direct imaging with deep learning. ECML PKDD, Springer, 2020. [paper]

Skills

LANGUAGES

5+ years of experience:
Python • Bash (awk, sed...) • LaTeX
3 years of experience:
C++ • Fortran

ML FRAMEWORKS

PyTorch • TensorFlow • scikit-learn • HuggingFace

HPC LIBRARIES

OpenMP • MPI • SYCL • Docker • Spack

Education

PhD Physics University of Cambridge

2013-2017 | Completed 2020
Studied PDE solvers to simulate wave propagation at material interfaces.

MPhil. Scientific Computing University of Cambridge

2012-2013 | Distinction

MSc Advanced Computer Science

University of Manchester
2010-2012 | Distinction

MEng Chemical Engineering and Chemistry University of Sheffield

2006-2010 | 1st Class

Links

https://github.com/ma595
https://linkedin.com/in/mattarcher595

Conferences & workshops

- Presenting "FTorch library for leveraging PyTorch in Fortran" at PyTorch Con 2025
- Organised and presented workshop titled "Enhancing the symbiosis between HPC and RSE communities" (RSEHPC@ISC24).

Teaching

- Taught "Introduction to Neural Networks with PyTorch" at the ICCS Summer School 2024 & 2025.
- Taught Deep Generative Models with TensorFlow 2 as part of SISSA Master's in HPC and Data Science (2021)