

Ian Wu

+44 7904777312 | ianyhwu.97@gmail.com | [linkedin.com/in/ianyhwu](https://www.linkedin.com/in/ianyhwu) | UK & US Dual Citizen

EDUCATION

Carnegie Mellon University

PhD Machine Learning

- **Advisor:** Aviral Kumar

Aug 2025 – Present

Pittsburgh, US

University College London

MSc Machine Learning

- **Grade:** Distinction, Dean's List (top 5% of cohort)
- **Thesis:** "Demonstrations-Involving Proximal Policy Optimisation". Advised by Minqi Jiang and Tim Rocktäschel.

Sep 2020 – Sep 2021

London, UK

University of Cambridge

BA, MSci Natural Sciences (Theoretical Chemistry)

- **Grade:** First Class Honours
- **Academic Awards:** Bundy Scholarship, B.C. Saunders Prize, Hong Kong Scholarship for Excellence Scheme

Oct 2016 – Jun 2020

Cambridge, UK

EXPERIENCE

Senior Data Scientist

C3 AI

- Utilised reinforcement learning with execution feedback to improve the reasoning and self-correction capabilities of C3 DSL code generation LLMs.
- Developed a novel inference-time algorithm that improves the performance of both open-source and proprietary LLMs on instruction-following tasks by leveraging an LLM judge for supervision during generation. Currently used across various applications to improve inference performance.
- Developed a novel synthetic data generation technique for scalably creating realistic domain-, modality- and style-specific multimodal RAG evaluation datasets. Currently used by customer-facing teams for benchmarking.

Apr 2023 – Jul 2025

London, UK

Data Scientist

C3 AI

- Trained 70B parameter LLMs for C3 DSL code generation. Scraped pre-training data, crowdsourced and generated finetuning data. Further aligned models to user preferences using DPO. These LLMs now power a chatbot application and code completion service used by hundreds of employees and customers.
- Implemented and maintained in-house LLM training infrastructure, based on DeepSpeed and Hugging Face Accelerate, capable of training (SFT, RLHF, DPO) models of up to 70B parameters by leveraging multi-node training.
- Finetuned open-source LLMs for question answering, significantly improving performance on in-house reading comprehension and RAG benchmarks. These models are used for RAG in several air-gapped production applications.

Sep 2021 – Apr 2023

London, UK

OTHER

Competitive Debater

Cambridge Union

- Competitive debater representing the Cambridge Union. Competed at over 30 national and international competitions. Three-time quarter-finalist and top-20 debater at the World Universities Debating Championships.
- Winner of the European Universities Debating Championships (2019), Europe's biggest debating competition at the university level, featuring over 100 universities.

Oct 2016 – Jan 2020

Cambridge, UK

Competitive Debater

Hong Kong National Debating Team

- Competitive debater representing the Hong Kong National Debating Team. Competed at various school-level international debating competitions.
- Semi-finalist and fourth best debater at the World Schools Debating Championships (2016), the world's biggest debating competition at the schools level, featuring over 60 national teams.

Sep 2014 – Aug 2016

Hong Kong

SKILLS

Programming Languages: Python

Developer Tools and Libraries: PyTorch, DeepSpeed, Hugging Face, FSDP, Git, Google Cloud Platform