Cemlyn Waters

cemlynwaters@gmail.com | +44 (0)7775 449333 | linkedin | github

Summary

Machine Learning Engineer, Senior Associate, at JPMorganChase, London. Ex-InstaDeeper who was building DeepPCB. Distinction in MSc Artificial Intelligence from Imperial College and First Class BSc in Mathematics from the University of Southampton.

Machine Learning Engineer at JPMorganChase — *December 2024 – Present*

Applied Machine Learning Engineer, Senior Associate, developing AI-powered products that automate processes, prioritise tasks, and improve decision-making. Responsibilities:

- Build and deploy scalable Machine Learning services integrated with strategic systems.
- Advance state-of-the-art AI applications in financial services, leveraging research in NLP, Computer Vision, and statistical machine learning.
- Communicate AI capabilities and results effectively to technical and non-technical audiences.

InstaDeep — *August 2022 – November 2024*

Research Engineer II — *May 2024 – November 2024*
Machine Learning Engineer II — *May 2023 – May 2024*
Machine Learning Engineer I — *August 2022 – May 2023*

Started in Research as a Machine Learning Engineer, later transitioned to Applied as a Research Engineer, building DeepPCB. Deploying software on distributed hardware with A100s. Accomplishments using Python and C++ with JAX, TensorFlow, Ray and DeepMind libraries:

- Reinforcement Learning (RL) distributed training library used internally for research and products, using Ray, JAX, TensorFlow and DeepMind libraries.
- A multi-phase pipeline, using Behavioural Cloning, Curriculum Learning and Search.
- Promoted as a result of achieving for all future experiments, 95% CPU and 39% GPU cost savings whilst reducing training time by 61%, for one of the applied revenue-making products. Primarily attained by implementing a distributed Asynchronous Proximal Policy Optimisation (APPO).
- Increased throughput by 321% and reduced all experiment costs by 60% by implementing a C++ gRPC server hosting GPU TensorFlow functions called by CPU only machines.
- Optimised asymptotic performance of multi-node distributed search by 300%, after profiling with Grafana & py-spy.
- Improved GPU utilisation whilst speeding up APPO learning by 400%, by interleaving batches.

Software Engineer at Levelise — *April 2021 – August 2022*

Levelise provide stabilisation services to the National Grid, operating in the UK and Australia, by coordinating the actions of a fleet of domestic batteries. Responsible for enhancing their cloud service, I accomplished this by:

- Optimising MySQL database queries, reducing data aggregation calculations run-time by 10x.
- Calculating the daily statistics for their IoT devices, deployed using AWS. Achieving 15x faster run-time.
- Automate fault detection and manage a fleet of customer battery systems - increasing detection coverage by 80%.
- Developing REST APIs, in Python and Go.

Education

- **MSc Artificial Intelligence** — Imperial College London
  Distinction — *Oct. 2019 – Sept. 2020*
- **BSc Mathematics** — University of Southampton
  First Class Honours — *Sept. 2016 – July 2019*

Publication

- Clément Bonnet *et al.* (2023). "Jumanji: a Diverse Suite of Scalable Reinforcement Learning Environments in JAX". Submitted: *NeurIPS 2023 Datasets and Benchmarks Track*. doi:10.48550/arXiv.2306.09884.
- Vosylius, Wang, **Waters** *et al.* (2020). "Geometric Deep Learning for Post-Menstrual Age Prediction Based on the Neonatal White Matter Cortical Surface". In: *Proceedings of UNSURE 2020*, pp. 174–186. Lecture Notes in Computer Science, Springer. doi:10.1007/978-3-030-60365-6_17.