Dylan Forde

Cork, Ireland

+353896013824 forde.dylan@gmail.com github.com/dylanbforde linkedin.com/in/dylanbforde

Research Projects

Block-Wise Differentiable Sinkhorn Attention (NeurIPS Manuscript)

- Designed a memory-efficient Optimal Transport kernel for ultra-long context alignment (16K+) on TPU hardware, achieving $O(L)$ HBM scaling versus the $O(L^2)$ footprint of standard attention methods.
- Derived a novel multi-transport-plan backward pass using four distinct matrices ($P^{(2,2)}$, $P^{(2,1)}$, $P^{(1,1)}$, $P^{(1,0)}$), eliminating divergent gradient bias inherent in standard implicit differentiation.
- Implemented high-performance JAX/Pallas kernels with 8-channel VMEM packing, Cody-Waite range reduction, and configurable banded Sinkhorn ($O(L \cdot W)$) optimised for TPU v6e (Trillium).

Intrinsic Quantification of Domain Shift Magnitude in Deep RL Agents

- Developed methods to quantify the effects of domain shift on RL agents using reactive exploration, evaluated across multiple OpenAI Gym environments. Results and code available on GitHub.

Multi-Omics Analysis for Hepatic Liver Cancer

- Integrating genomic, transcriptomic, and proteomic data to identify predictive biomarkers of liver cancer progression, applying deep learning and statistical modeling to multi-omics pipelines.

Experience

Valeo Vision Systems — **Galway, Ireland**
Deep Learning Engineer — *September 2025 – Present*

- Developed a high-efficiency Generative AI pipeline processing 10,000+ parking video traces, achieving 95% Model Flops Utilization (MFU) via complex object addition/removal using advanced segmentation, masking, and inpainting.
- Spearheaded deep learning research on fully generated images, employing SSIM, heatmaps, and noise analysis to evaluate the impact of crops, augmentations, and background replacements on downstream model performance.
- Engineered a hardware-agnostic Height Map Processing optimization in JAX, cutting trace processing time by 89% (45 to 5 minutes); previously delivered a 7$\times$ speedup to 3D pointcloud rasterization via JAX/Pallas kernels.
- Investigated and proposed agentic workflows for CI/CD automation, conducting feasibility studies on local versus throttled AI agents and using agents to produce comprehensive codebase documentation from scratch.
- Developing end-to-end autonomous driving algorithms spanning data pipelines through model training.

Microsoft Research — **Dublin, Ireland**
Software Engineering Intern — *June 2024 – September 2024*

- Developed state and API mocking solutions for production and testing environments, enabling rapid preview of CMS-driven live changes and saving significant engineering hours.

Microsoft Research — **Dublin, Ireland**
Software Engineering Intern — *March 2023 – September 2023*

- Led an initiative to evaluate Copilot outputs using NLP, semantic accuracy metrics, and topic modelling.
- Hackathon: fine-tuned an LLM as an onboarding assistant for new hires; conceptual platform addressing M365 educational suite challenges.

Agnicio — **Rotterdam, Netherlands**
Data Science & Machine Learning Intern — *June 2022 – September 2022*

- Benchmarked proprietary forecasting models against top Kaggle baselines; developed new LightGBM, XGBoost, and ARIMA models. Transitioned proprietary forecasting models to a SaaS approach.

Technical Projects

Personalized Cancer Vaccine Design with Graph AI — Built an end-to-end pipeline from TCGA mutations to peptide candidates; modeled peptide–MHC binding with GNNs trained on IEDB, validated with UniProt context, ranked neoantigens by predicted affinity.

Cardiac MRI Segmentation (3D SegFormer) — Implemented a 3D MixTransformer backbone with overlapping patch embeddings, stochastic depth, and MLP decoder for multi-scale fusion; voxel-wise masks via trilinear upsampling.

MRNet MRI Classification (ViT + Slice Attention) — Tokenized 2D slices into patch sequences with multi-head self-attention; aggregated CLS tokens via inter-slice attention and learned positional embeddings for volume-aware diagnosis.

ODE Modelling for Drug Response Prediction — Analyzed GRN outputs (GENIE3, PPCOR) to derive hub genes and consensus interactions; integrated features into downstream survival/response models.

Deep RL Agents for OpenAI Gym/Atari — Implemented agents under domain shifts with spaced-repetition-style training; benchmarked across classic control and Atari environments.

Education

University College Dublin — **Dublin, Ireland**
M.Sc. Artificial Intelligence for Medicine and Medical Research (**GPA:** 3.75/4.00) — *August 2024 – September 2025*

University College Cork — **Cork, Ireland**
B.Sc. Data Science and Analytics (**GPA:** 3.68/4.00) — *August 2020 – September 2024*

Skills

Languages: Python, R, SQL, TypeScript, C#, Rust, Java, Haskell, LaTeX
Deep Learning: PyTorch, JAX, JAX-Pallas, scikit-learn, NumPy, Pandas, PySpark
ML & AI: Deep Learning, Reinforcement Learning, Optimal Transport, Generative AI, Statistical Modeling, Computer Vision, Forecasting
Cloud & Hardware: Google Cloud (Vertex AI, TPU v6e), AWS, Azure, Docker, Kubernetes
Tools: Git, Linux (Nix, Arch, RedHat), DVC, MySQL, MongoDB

Additional

Google Workshops (2022, 2024, 2025), Amazon Mentoring Program (2022), KPMG Finance Internship (2019), CHANEL Finance Internship (2017)