



JAMES M. SHORE
Address: [REDACTED] Hemlock
Shoreville
EPOC

[REDACTED]

Course	University	Dates	Grade/%
Artificial Intelligence M.S.c	King's College London	2022 to 2023	Merit/(62.5%*)
Computer Science B.S.c	Newcastle University	2018 to 2021	1st Class/(71.5)%

*Current average, final grades released in November

SUMMARY STATEMENT

- Impactful Machine Learning Engineer, with a 3-year record in ML and 9-years programming. Highly proficient in Python, PyTorch and SQL. Discovered method to improve drug discovery by 45.96%. Expertise in Python data structures, optimising scalable models with DistributedDataParallel, fluent with torch.nn and with high-performance computing using C++ and CUDA

RESEARCH

- **Novel Application of Large Language Models (LLMs) for In Silico Drug Discovery: State-of-the-Art** 09/2022 to Present
 - Engineered deep learning architectures using an advanced integration of chemical and proteomic molecular data, utilising RDKit and BioPython for cheminformatic feature engineering, implementing cross-attention in torch.nn. Enhanced Pearson R, 0.485 to 0.7079, outperforming all prior benchmarks
 - Utilised Distributed Data Parallel (DDP) for efficient gradient averaging, effective training of 1.2 billion model parameters across 4x A100 GPUs, realised a 60-fold reduction in training time
 - Released findings to r/machinelearning, amassing 50,000+ views in 3 hours, sparking global academic interest, recognition from the KCL AI Director Prof. Ventre, and offers to co-author with top-tier professors from TU Munich
- **Edge Detection in Computer Vision via Optimised Image Preprocessing and Canny Algorithms** 09/2022 to 08/2023
 - Engineered edge detection on noisy images, achieving human results with only 12 training images, F1-score of 0.711
 - Developed adaptive luminance scaling and weighted RGB-to-grayscale conversion, in MATLAB R2022b
 - Built guided filtering and Gaussian smoothing algorithms, enhancing edge discriminability, coded a Canny edge-detection algorithm with hyperparameter grid-search. Achieving an 83% mark (top 1%)
- **GPU-Accelerated Deep Learning for Time-Series Signal Classification in Physical Activity Prediction** 09/2020 to 06/2021
 - Engineered a data pipeline capable of ingesting and preprocessing 500,000 multivariate time-series signals, employing feature extraction techniques that eliminated the need for manual annotation
 - Developed a C++ algorithm with cuFFT to generate frequency-domain spectrograms and a heuristic feature selection using information theory, accelerating processing speed from 180s to 20s
 - Designed a domain-specific deep learning model that integrates 1x5 CNN kernels for optimised frequency classification, achieving an 18% accuracy boost compared to pre-trained architectures such as ResNet, MobileNet and VGG19. Code-base adhered to software engineering best practices, DOCSTRINGS etc., received a first-class mark
- **Neuronal Genomic Characterisation of Cyclin-Dependent Kinase 5 in C. Elegans Neurobiology (->)** 01/2021 to 03/2021
 - Mastered computational genomic analysis of CDK5, a key regulator in neuronal signaling
 - Employed sequence alignment pipeline: gene prediction of FASTA sequences with Ensembl gene summary algorithm, optimised BLAST search for FGENSESH similarity, EBI Clustal-Omega to generate Phylogenetic trees, then performed 3D protein structure prediction using PSIPRED algorithm
 - Identified ATP binding site at position 31 which is not currently annotated on SwissProt

EXPERIENCE

- **Lab Leader | KCL AI Society** 09/2022 to 04/2023
 - Leadership and coordination of KCL AI Kaggle Competition