

Kieran Didi

EDUCATION

MPhil Computational Biology (Appl. Maths), Cambridge University, St. John's College 10/2022 – 09/2023

- Lectures include adv. graph ML, protein design, information theory, probabilistic ML
- Supervising [algorithms course](#) for CS students (graphs, advanced data structures)

M.Sc. Biochemistry (focus CS), Ruprecht-Karls-University Heidelberg 10/2021 – Present

- Lectures from CS/Maths/Physics faculty, including machine learning I+II, GPU computing, operating systems, networks, software development, algorithms
- Lectures from Biology/Chemistry faculty, including bioinformatics, multi-omics analysis, simulation methods, journal club ML for biology
- Designed and lectured a course on [data analysis with Python](#)

Cambridge University, St. John's College (scholarship exchange year) 09/2020 – 10/2021

Natural Sciences Part II, lectures including math. methods, symmetry, cheminformatics, theoretical chemistry, scientific programming

B.Sc. Biochemistry, Ruprecht-Karls-University Heidelberg 10/2018 – 10/2021

Grade 1.0 (highest grade of 73 students), additional lectures in macroeconomics, econometrics, drug discovery and statistical learning

WORK EXPERIENCE

EMBL-EBI Cambridge, UK
Machine Learning Engineer Intern 10/2022 – Present

- Critically test and evaluate BioML models (preprint in preparation)
- Currently implementing protein language models inspired by BERT, GPT and T5

CSIRO Sydney, team Translational Bioinformatics Sydney, Australia
Software Engineering Intern (Cloud-native genomics) 07/2022 – 10/2022

- Developing [cloud-native Python software](#) for omics analysis (paper under review)
- Presented keynote at [eSCAMPS symposium](#) and at [ABACBS conference](#)
- Published [open-source Terraform module](#) for Spark on AWS (>400 downloads)

PTNG Consulting Melbourne, Australia
Consultant for ML/bioinformatics projects 06/2022 – Present

- Deliver and communicate insights via literature research/ML and bioinformatics analysis, also implement custom algorithms
- Example topics: [protein design](#) and structural analysis of antibodies

Chemistry Department, Cambridge University Cambridge, UK
Thesis Student, Bernardes/Knowles/Sormanni Lab 09/2020 – 10/2021

- Designed peptide therapeutics and developed protein screening system (published in [Nature Communications](#))
- Enabled improved classification results by building large-scale database for high-quality antibody sequences, incl. quality control
- Developed ML models and protein sequence processing pipeline to quantify nativeness of antibody sequences (PyTorch)

BioMed X Innovation Center (with Janssen Pharmaceuticals) Heidelberg, Germany
Research Intern 08/2019 – 04/2020

- Established a screening system for peptides in autoimmune diseases for multi-national pharmaceutical company, throughput improvements enabling personalized patient screens
- Developed and presented novel approach based on antibody fragments to BioMedX founder, board members and staff (50)

SKILLS

Programming: Python, R, Java, C++ (prior experience)

Machine Learning: PyTorch, PyTorch Lightning, W&B, Hydra, sklearn, JAX (prior experience)

Cloud Computing/HPC: Docker, CUDA, AWS, Terraform (Infrastructure-as-code), Spark, SLURM/PBS

Web Development: HTML/CSS, JavaScript, MERN stack, focus on backend

Protein Engineering: PyMol, ChimeraX, BioLuminate, ML Tools, biopython, mmseqs2, RDKit, basic Rosetta

Teaching: Designed and held undergraduate lectures on [data science with Python](#), tutored maths/chemistry/biochemistry, supervise [algorithms course](#), publish [blog](#)

Open Source Development: Implement biological oracle in [DNA Diffusion Project](#) of [OpenBioML Lab](#), published [Terraform module](#) for [biological sequence processing](#), work on [multimodal healthcare project](#)

Online Coursework: Algorithms I+II (Princeton), Deep Learning (Andrew Ng), ML with Graphs (Stanford)

ML Summer Schools: OXML, EEML (presented [GNN paper](#)), SMLW, DLAI6, Medical DL, Resource-aware ML

CERTIFICATES

- 2022: AWS Certified Cloud Practitioner
- 2022: Web Development BootCamp TechLabs (6 months)
- 2022: NVIDIA: Fundamentals of Deep Learning with Multiple GPUs
- 2021: Cambridge i-Teams certificate for consulting project at biotech startup (ADC technology)
- 2019: Data Science BootCamp TechLabs (6 months)

LEADERSHIP AND AWARDS

- 2023: [Software Sustainability Fellowship](#)
- 2023: [Polaris Fellowship](#)
- 2022: [AI Fellowship](#) by Hummingbird VC
- 2022: [Kurt Hahn Scholarship](#) at Cambridge University (academic merit)
- 2022: Scholarship from both DAAD and Studienstiftung for master's studies at Cambridge
- 2021: [Sartorius scholarship](#) for master's studies at Heidelberg (academic merit)
- 2021: [Marsilius Certificate](#) for statistical learning: econometric predictions via LSTMs
- 2020: Scholarship for exchange year at Cambridge University
- 2019: Digital Shaper Award for project at TechLabs program ([water quality predictions](#))
- 2018: Scholarship of German Academic Foundation based on intellectual ability
- 2018: Biology Olympiad, [2nd place in Germany](#) (>2,000 participants), silver medal at the International Competition in Tehran (best 30% of global selection of talents)

VOLUNTARY EXPERIENCE

Nucleate UK Cambridge, UK
Communications Lead, Cambridge Chapter 07/2022 – Present

- Organized nationwide comms strategy for empowering biotech talents

German Biology Olympiad Association Kiel, Germany
Steering committee 05/2019 – 07/2022

- Co-ordinated selections rounds: Coached 45 participants/year
- Enabled selected participants exposure to leading researchers via internships at prestigious institutions (e.g. Max Planck Institute) each year, launched new formats (e.g. participation in summer schools)

Student parliament Heidelberg Heidelberg, Germany
Conference administration 11/2018 – 09/2020

- Designed, chaired and summarized biweekly debates
- Reduced debating time by >1 hour/debate by optimising procedures

INTERESTS

- Handball player, 3x state championship winners, first local team to reach statewide league
- Authored a fantasy book about a teenager able to stop the time (220 pages)
- Published opinions and reports (10) in newspapers (e.g. Kölner Stadt-Anzeiger)

- Science communication at preLights, [publishing highlights](#) of new “ML for biology” preprints
- Enthusiastic guitar player for 12 years, teamed up with drummer for cover songs