dennis akar

denizhanak@gmail.com | denden.dev | github.com/dendenakar | UK Citizen

EDUCATION

University of Cambridge 2021 - 2022
MPhil in Advanced Computer Science
- Pass with distinction with 81.20% GPA.
- Awarded the £5,000 ACS MPhil Scholarship for academic excellence.
- Researched geometric DL for molecular graphs (drug discovery) supervised by Prof Pietro Liò & Dr Cristian Bodnar.

University of Manchester 2018 - 2021
BSc Computer Science and Mathematics
- First Class Honours with 83.36% GPA.
- Awarded Certificate of Excellence for top 10% graduating students.

EXPERIENCE

MATS: Foundations of Mechinterp (Lee Sharkey) - Researcher May 2023 - Jan 2024
- Investigating "Attention Head Superposition" in language models with Chris Mathwin and Lee Sharkey.
- Discovered how certain properties of the model affect model performance.
- Facilitated Alignment 201 reading group for 5 MATS scholars.

ARENA June 2023
- Self-studied Redwood Research's MLAB curriculum to work as a TA (for the mechanistic interpretability chapter) and as a participant (for the training LLMs at scale chapter) for ARENA (Alignment Research Engineer Accelerator).
- Aided participants on examining and interpreteing Indirect Object Identification, balanced bracket classification, superposition, and OthelloGPT.

MATS: Mechanistic Interpretability (Neel Nanda) - Researcher Nov 2022 - Jan 2023
- Applied the original and extended logit lens to the IOI task across a set of GPT-2 sized models (extended DLA). Extended logit lens uses consecutive layers at the end of the model to map the residual stream to logit space.
- Found the tendency for certain models (e.g. GPT-Neo) to "flip" i.e. assign an *extremely low probability* throughout the model to the token that it will eventually output and used extended DLA to analyze how this tendency changes.

CancerAI (University of Cambridge) - Research Assistant Jul 2022 - Oct 2022
- Researched explainable AI for use by clinical oncologists using **Tensorflow** and **PyTorch**.
- Developed front-end for VIIDA, an application for analyzing, modelling, explaining, and predicting cancer-related data with **Flask** and **React**.

Cambridge Cancer Genomics - Software Engineer Intern Jun 2019 - Sep 2019
- Integrated features and fixed bugs for the precision oncology platform OncOS backend using **Python** and **Flask**.
- Built a **full-stack** internal monitoring system for OncOS infrastructure to manage genomic data and processes.
- Researched variational autoencoder algorithms related to DNA sequence compression for SomaticNET, a neural network for evaluating tumor variants, using **Tensorflow (Python)**, **Bash**, **pysam** and **Annoy**.

PROJECTS

Geometric CW Networks - MPhil Thesis 2021 - 2022
- Introduced geometric inductive priors [E(3) invariance and equivariance] to a GNN with a topological inductive prior, in this case CW Networks (CWNs), an architecture in which graphs are "lifted" into higher order hypergraphs using CW complexes, to construct Geometric CW Networks (GCWNs).
- Used **PyTorch**, **PyTorch Geometric**, **gudhi**.

Topological Neural Processes - BSc Thesis 2020 - 2021
- Built a novel machine learning model for extracting latent information of topological structures of input (topological data analysis) for Conditional Neural Processes (a neural model which meta-learns a stochastic process) using **Tensorflow**, **matplotlib**, **pickle**, **gudhi**, and **numpy**; supervised by Dr Tingting Mu and Dr Cristian Bodnar.

LEADERSHIP

AI Safety Facilitator and Tutor 2023
- Facilitating AI Safety Fundamentals 101 and 201 reading groups and tutoring for MLAB/ARENA (including my modifications for Anki support) for MATS research fellows and graduates from the University of Cambridge.