XIANGYU ZHAO

82 The Belvedere, Homerton Street, Cambridge, United Kingdom CB2 0NU
Telephone: +44 7422 900440 (UK) ◇ +86 185 1241 5307 (China)
Email: xz398@cantab.ac.uk ◇ Website: https://victorzxy.github.io/ ◇ GitHub: VictorZXY

EDUCATION

PhD in Electrical and Electronic Engineering *January 2023 – Present*
Imperial College London, United Kingdom
Research topic: self-supervised learning on complex hypergraphs via adversarial collaboration
Supervisor: Dr Yiren Zhao
Expected graduation: January 2027

BA & MEng (Hons) in Computer Science *October 2018 – June 2022*
Trinity College, University of Cambridge, United Kingdom
BA final year result: 2.i (Overall: 70.5%; Units: 86%; Dissertation: 81%)
MEng result: Distinction (Overall: 81.2%; Modules: 79.7%; Dissertation: 83%; Ranking: 7/25)
Awards: Senior Scholarship; Exam Prize 2022

WORK EXPERIENCE

Department of Computer Science and Technology, University of Cambridge *January 2023 – Present*
Visitor *Cambridge, United Kingdom*

- Undertake research work for the PhD in EEE degree at Imperial College London;
- Cambridge co-supervisor: Prof Pietro Liò.

Department of Computer Science and Technology, University of Cambridge *June – September 2022*
Undergraduate Research Opportunities Programme (UROP) – Research Intern *Cambridge, United Kingdom*

- Research project: long and short-range attentions for complex hypergraph data;
- Supervisors: Dr Yiren Zhao, Prof Robert Mullins.

ByteDance *June – September 2021*
Algorithm Engineer Intern *Beijing, China*

- Trained a model based on causal forest and double ML for TikTok Lite's user growth campaign;
- Increased per-user successful invitation count by 43.9% while decreasing its cost by 56.1%.

PUBLICATIONS

- **Xiangyu Zhao** and Sean B. Holden (2022). Towards a Competitive 3-Player Mahjong AI using Deep Reinforcement Learning. In *2022 IEEE Conference on Games (CoG)*.
- **Xiangyu Zhao** and Sean B. Holden (2022). Building a 3-Player Mahjong AI using Deep Reinforcement Learning. *arXiv preprint arXiv:2202.12847*.

PROJECT EXPERIENCE

Long and Short-Range Attentions for Complex Hypergraph Data *June 2022 – Present*
UROP Project

- Build benchmark hypergraph datasets from large-scale real-world data;
- Design new attention mechanisms for GNNs to operate on hypergraph data.

Task-Agnostic Graph Neural Network Evaluation via Adversarial Collaboration *November 2021 – May 2022*
MEng Dissertation

- Developed a conceptually novel, principled, task-agnostic, and stable framework for evaluating GNNs through contrastive self-supervision, without needing handcrafted augmentations;
- Distinguish GNNs of different expressivity by having them directly compete against each other on unlabelled graphs, and ensures that more expressive GNNs can win by producing more informative graph embeddings.
- Source code: https://github.com/VictorZXY/GraphAC

Improving Graph Generative Models via Expressive Graph Neural Networks *March – April 2022*
MEng Representation Learning on Graphs and Networks Mini-Project

- Investigated the expressivity of GNNs under the context of molecular graph generation;
- Replaced R-GCN in GCPN with GIN, PNA and GSN, and significantly improved GCPN's performance.
- Source code: `https://github.com/VictorZXY/expressive-graph-gen`

Multi-Agent Deep Q-Learning for the Berry Poisoning Game *February – March 2022*
MEng Advanced Topics in Machine Learning Coursework (Reinforcement Learning)

- Performed a multi-agent DQN method on the Berry Poisoning Games, and investigated on the agent performance with respect to different game environment parameters;
- Successfully trained agents for the game, with high transferability across different game environment parameters.
- Source code: `https://github.com/VictorZXY/dqn-berry-poisoning`

Function Autoencoders: A Neural Network Approach to Gaussian Processes *November 2021 – January 2022*
MEng Probabilistic Machine Learning Investigative Project

- Investigated a neural network alternative to Gaussian processes, and introduced the function autoencoders that preserve GPs' own advantages and avoid their weaknesses with NNs' benefits;
- Models successfully learnt distributions over random functions, and performed decently on a 1-d regression task.
- Source code: `https://github.com/VictorZXY/function-autoencoder`

Building a Simulator and Emulator for Traffic Signaling *November 2021 – January 2022*
MEng Machine Learning and the Physical World Group Project

- Carried out simulation of an urban traffic signalling system, and built an emulator to search for optimal traffic signal scheduling using Bayesian optimisation;
- Introduced multiple scheduling schemes to overcome the problem of exploding search space, without sacrificing flexibility or descriptiveness.
- Source code: `https://github.com/VictorZXY/traffic-sim`

A Neural Network Approach to Named Entity Recognition on Noisy User-Generated Texts
MEng Natural Language Processing Final Assignment *November – December 2021*

- Investigated a bidirectional LSTM structure for named-entity recognition on social media texts, and explored various data-processing techniques in order to improve the model's performance;
- Model trained with data-processing techniques applied achieved significant improvements on prediction accuracy on the W-NUT 2017 shared task, compared to the same model trained without any data optimisation.
- Source code: `https://github.com/VictorZXY/nlp-assignments-ner`

Deep Reinforcement Learning for Mahjong *October 2020 – May 2021*
BA Dissertation

- Designed an informative and compact data structure for encoding observable Mahjong states, built an AI for Mahjong by pre-training CNN models for each action, and improved the major action's model with self-play RL;
- Models achieved near state-of-the-art level, and RL significantly improved the agent's win rate from SL;
- Paper presented at the 2022 IEEE Conference on Games (CoG).
- Source code: `https://github.com/VictorZXY/meowjong`
  Related publication: `https://arxiv.org/abs/2202.12847`

Investigating Adversarial Examples for Deep Residual Networks *February – March 2021*
Third Year Deep Neural Networks Mini-Project

- Investigated on targeted/black-box adversarial attacks on ResNets using one-shot/iterative FGSM algorithms, and experimented the transferability of each FGSM attack to other image inputs and neural networks.
- Source code: `https://github.com/VictorZXY/dnn-assignments`

An Analysis of Clickstream Data for Online Shopping *November – December 2020*
Third Year Data Science Final Practical

- Analysed a clickstream dataset on an online shopping website, trained various ML models to predict the costumers' potential willingness to pay a premium price, and identified the most important features for ML prediction.
- Source code: `https://github.com/VictorZXY/datasci-pnp-final-practical`

Trading Assistant for IMC *January – March 2020*
Second Year Group Project

- This is a web service which holds market data and responds to queries on demand in a human-like manner, by automating the trader side of the process using speech recognition and NLP;
- Responsible for writing speech-to-text and text-to-speech libraries;
- Efficiently built the libraries with high accuracies, and provided clean interfaces that were used conveniently by the NLP and back-end teammates, improving the system's overall performance.

AWARDS

Google Hash Code	*2020*	Team ranked 1331/10724 globally
LeetCode Programming Contest (LCCUP)	*2020*	Team ranked 127/2575
UK & Ireland Programming Contest (UKIEPC)	*2019*	Team ranked 54/191
British Informatics Olympiad (BIO)	*2017*	National finalist (top 15)
British Physics Olympiad (BPhO)	*2017*	Gold Medal
British Chemistry Olympiad (BChO)	*2017*	Gold Medal
British Mathematics Olympiad (BMO)	*2016, 2017*	Distinction
Senior Team Mathematics Challenge (STMC)	*2017*	Team won regional 2nd place
Senior Mathematics Challenge (SMC)	*2016, 2017*	Gold Medal (Full marks)
American Mathematics Contest (AMC12)	*2016*	Distinction (Global top 1%)
American Mathematics Contest (AMC10)	*2015*	Distinction (Global top 1%)
National Olympiad in Informatics (NOI)	*2014*	2nd Place

SKILLS

Languages	Chinese (Native), English (Fluent, IELTS Band 8.0)
Programming	Python, Java, C, C++, C#, OCaml, StandardML, Prolog, SQL, Cypher, Verilog, LaTeX, Swift
ML Libraries	PyTorch, TensorFlow, Keras, Scikit-learn, PyG, DGL, TorchDrug, GPy, GPyOpt, EconML
Music	Clarinet (Grade 10, soloist of the Cambridge University Chinese Orchestra Society)
Sports	Kendo (University team), Football (High school team), Badminton