# XIANGYU ZHAO

48 Haydn Tower, 50 Wandsworth Road, London SW8 2FN
Personal Website: https://victorzxy.github.io/
Email: x.zhao22@imperial.ac.uk ◇ GitHub: VictorZXY

## EDUCATION

**PhD in Electrical and Electronic Engineering** *January 2023 – Present*
***Imperial College London, United Kingdom***
Research topic: Representation Learning on Higher-Order Graphs
Supervisor: Dr Yiren Zhao
Award: Electrical and Electronic Engineering PhD Scholarship
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** *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

- Xiandong Zou, **Xiangyu Zhao**, Pietro Liò, and Yiren Zhao (2023). Will More Expressive Graph Neural Networks do Better on Generative Tasks? *arXiv preprint arXiv:2308.11978*. Also under review at *The 2nd Learning on Graphs Conference (LoG 2023)*.
- **Xiangyu Zhao**, Zehui Li, Mingzhu Shen, Guy-Bart Stan, Pietro Liò, and Yiren Zhao (2023). Semi-HyperGraph Benchmark: Enhancing Flexibility of Hypergraph Learning with Datasets and Benchmarks. Under review at *The 11th International Conference on Learning Representations (ICLR 2024)*.
- **Xiangyu Zhao**, Zehui Li, Mingzhu Shen, Guy-Bart Stan, Pietro Liò, and Yiren Zhao (2023). Hybrid Graph: A Unified Graph Representation with Datasets and Benchmarks for Complex Graphs. *arXiv preprint arXiv:2306.05108*.
- **Xiangyu Zhao**, Hannes Stärk, Dominique Beaini, Yiren Zhao, and Pietro Liò (2023). Task-Agnostic Graph Neural Network Evaluation via Adversarial Collaboration. *The 11th International Conference on Learning Representations (ICLR 2023) Machine Learning for Drug Discovery (MLDD) Workshop*.
- **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

**Investigating GNN Expressiveness in Graph Generation Tasks** *2023*

- Improved GNN-based graph generative models with more advanced GNNs;
- Showed that GNN expressiveness in graph prediction does not correlate to its performance in graph generation.
- Source code: https://github.com/Yqcca/graph-generative-models
  Publication: https://arxiv.org/abs/2308.11978

**Unifying Higher-Order Graph Representation with New Datasets and Benchmarks** *2023*

- Introduced the concept of hybrid graphs, a unified definition for higher-order graphs;
- Developed a collection of hybrid graph datasets with an extensible evaluation framework.
- Project homepage: `https://zehui127.github.io/hybrid-graph-benchmark/`
  Publication: `https://arxiv.org/abs/2306.05108`

**Task-Agnostic Graph Neural Network Evaluation via Adversarial Collaboration** *2022*

- 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 expressiveness 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.
- Paper published at the ICLR 2023 Machine Learning for Drug Discovery (MLDD) Workshop.
- Source code: `https://github.com/VictorZXY/GraphAC`
  Publication: `https://arxiv.org/abs/2301.11517`

**Multi-Agent Deep Q-Learning for the Berry Poisoning Game** *2022*

- 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** *2022*

- 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** *2022*

- 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 optimise 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** *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** *2021*

- 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 published at the 2022 IEEE Conference on Games (CoG).
- Source code: `https://github.com/VictorZXY/meowjong`
  Publication: `https://ieeexplore.ieee.org/document/9893576`

**Investigating Adversarial Examples for Deep Residual Networks** *2021*

- 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** *2020*

- Analysed a clickstream dataset on an online shopping website, trained various ML models to predict the users' 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`

## TEACHING EXPERIENCE

**Department of Electrical and Electronic Engineering, Imperial College London** *January 2023 – Present*
*Teaching Assistant & Project Supervisor* *London, United Kingdom*

· Projects: Investigating GNN Expressiveness in Graph Generative Tasks (UROP Project 2023), Neural Weather Forecasting with High-Order Graphs (MEng Final-Year Project 2023–2024).

**Department of Computer Science and Technology, University of Cambridge** *January 2023 – Present*
*Teaching Assistant & Course Supervisor* *Cambridge, United Kingdom*

· Courses: Algorithms 1 (2022–2024), Algorithms 2 (2023–2024), Foundations of Computer Science (2023–2024), Introduction to Probability (2022–2024).

## SKILLS

| | |
|---|---|
| **Languages** | Chinese (Native), English (Fluent, IELTS Band 8.0) |
| **Programming** | Python, Java, C, C++, C#, OCaml, StandardML, Prolog, SQL, LaTeX |
| **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 (High school team) |