XIANGYU ZHAO

82 The Belvedere, Homerton Street, Cambridge, United Kingdom CB2 0NU
Telephone: +44 7422 900440 (UK) ⋄ +86 185 1241 5307 (China)
Email: x.zhao22@imperial.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: principled self-supervised learning on complex 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 *January – July 2023*
Visiting Student *Cambridge, United Kingdom*

- Undertake research work for the PhD in EEE at Imperial College London, co-supervised by Prof Pietro Liò;
- Roles: Teaching Assistant for Algorithms 1, Supervisor for Introduction to Probability.

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**, 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. *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

Hybrid Graph: A Unified Graph Representation with Datasets and Benchmarks for Complex Graphs *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 accepted 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`

Improving Graph Generative Models via Expressive Graph Neural Networks *2022*

- Investigated the expressiveness 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 *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 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 *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`

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