

# XIANGYU ZHAO

Burrell's Field, Trinity College, Grange Road, Cambridge, United Kingdom CB3 9DH

Telephone: +44 7422 900440 (UK) ◊ +86 185 1241 5307 (China)

Email: xz398@cam.ac.uk ◊ Website: <https://victorzxy.github.io/> ◊ GitHub: VictorZXY

## EDUCATION

**Trinity College, University of Cambridge, United Kingdom**

October 2018 – Present

BA + MEng (Hons) in Computer Science, expected graduation time: June 2022

BA final year result: 2.i (Overall grade 70.5%; Units: DNNs 88%, Data Science 84%; Dissertation: 81%)

## WORK EXPERIENCE

**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 event;
- Increased per-user successful invitation count by 43.9% while decreasing its cost by 56.1%.

## PUBLICATION

- **Xiangyu Zhao** and Sean B. Holden. Building a 3-Player Mahjong AI using Deep Reinforcement Learning. *arXiv preprint arXiv:2202.12847*, 2022.

## PROJECT EXPERIENCE

**Learning with CLASS: Collaborative Learning with Adversarial Self-Supervision**

November 2021 – Present

MEng Dissertation

- Designed a self-supervised learning framework for GNNs using a novel loss function derived from Barlow Twins, which can efficiently rank different GNNs with respect to their expressivity
- Dissertation to be submitted by 27 May 2022

**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**

*BA Dissertation*

*October 2020 – May 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 submitted to IEEE Conference on Games 2022.
- Source code: <https://github.com/VictorZXY/meowjong>
- Related publication: <https://arxiv.org/abs/2202.12847>

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

*Third Year Deep Neural Networks Mini-Project*

*February – March 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**

*Third Year Data Science Final Practical*

*November – December 2020*

- 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**

*Second Year Group Project*

*January – March 2020*

- 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.

## **iSport**

*First Year Interaction Design Group Project*

*May 2019*

- This is an Android application that gives a daily weather forecast information on the most relevant weather based on the selection of sports for the athletes at the University of Cambridge;
- Responsible for layout design and interaction processing, and built an aesthetic, easy-to-use and efficient UI.
- Source code: <https://github.com/VictorZXY/iSport>

## **AWARDS**

<b>Google Hash Code</b>	2020	Team ranked 1331/10724 globally
<b>UK &amp; Ireland Programming Contest (UKIEPC)</b>	2019	Team ranked 54/191
<b>British Informatics Olympiad (BIO)</b>	2017	National finalist (top 15)
<b>British Physics Olympiad (BPhO)</b>	2017	Gold Medal
<b>British Chemistry Olympiad (BChO)</b>	2017	Gold Medal
<b>British Mathematics Olympiad (BMO)</b>	2016, 2017	Distinction
<b>American Mathematics Contest (AMC10/12)</b>	2015, 2016	Distinction (Global top 1%)
<b>National Olympiad in Informatics (NOI)</b>	2014	2nd Place

## **SKILLS**

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