Yuchen Zhu

London, UK
Tel: +447496803878
Email: yuchen.zhu.18@ucl.ac.uk
Website: yuchen-zhu.github.io

Research Statement

I am a PhD candidate passionate about developing responsible AI methodologies as well as theoretical frameworks to ensure its safe deployment, including but not limited to applications in medical and policy decision making. This led me to my projects focused on causal machine learning, causal abstraction, and large language model safety.

Education

2020 - Present   **PhD in Foundational Artificial Intelligence, University College London**
Supervisors: Dr. Matt Kusner, Prof. Ricardo Silva. ELLIS Student, host: Dr. Dominik Janzing.
Interests: causal inference and abstraction; alignment of large models.
Teaching assistant for Gatsby Unit courses on probabilistic learning (with Prof. Peter Orbanz) and on kernel methods (with Prof. Arthur Gretton).

2018 - 2020   **MSc in Machine Learning *(Distinction)*, University College London**
Topics covered: supervised learning, natural language processing, deep learning, reinforcement learning, probabilistic and unsupervised learning.

2015 - 2018   **BA, MA in Mathematics, University of Cambridge**
A broad range of courses are covered in linear algebra, probability and statistics, and pure maths.

Research Experience

01.2023 - 03.2023   **Research Intern, Microsoft Research Cambridge**
Host**:** Nick Pawlowski, Cheng Zhang
Topic**:** Scalable causal effect estimation.
Summary: Developed identification and estimation strategies of causal effects under many mismeasured treatments. **Project code integrated into the company code base.**

06.2022 - 12.2022   **Research Intern, Amazon Research Tuebingen**
Host**:** Dominik Janzing
Topic**:** Causal Aggregation of Micro-Variables.
Summary: Investigated confounding properties under aggregation of many micro-variables, a question at the fundamentals of Causal Representation Learning. **Research paper accepted at CLeaR 2024.**

10.2021 - 09.2020   **Industrial research collaboration for Master's thesis, Sana Labs**
Project title**:** Causal machine learning for educational recommender systems.
Summary**:** Suggested my own Master's research project, and obtained data from this EdTech startup. Developed theory and algorithms for estimating treatment effects under hidden confounding for online education. **Thesis grade: Distinction**

10.2015 - 06.2018   **Machine Learning Scientist Intern, Pace Revenue, London**
Summary of contribution**:** Implemented the company's first simulation of RL agents for pricing, in tensorflow. Built cloud parallel training in Docker, Kubernetes and Argo. Wrote unit tests.

10.2015 - 06.2018   **Summer Research Student, DAMTP, University of Cambridge**
Topic**:** Optimal Flow Estimation

Publications and Preprints

2024 (Ongoing)   **Unsupervised Causal Abstraction**

Short version @ NeurIPS workshop   **Yuchen Zhu***, Sergio Hernan Garrido Mejia*, Bernhard Schoelkopf, Michel Besserve

CLeaR 2024   **Meaningful Causal Aggregation and Paradoxical Confounding**

Yuchen Zhu, Kailash Budhathoki, Jonas Kuebler, Dominik Janzing

UAI 2022
(Oral, 5% accept. rate)   **Causal Inference Under Treatment Measurement Error: A Nonparametric Instrumental Variable Approach.**

Yuchen Zhu, Limor Gultchin, Arthur Gretton, Matt Kusner, Ricardo Silva

NeurIPS 2021   **Causal Effect Inference for Structured Treatment**

Jean Kaddour, **Yuchen Zhu**, Qi Liu, Matt Kusner, Ricardo Silva

ICML 2021   **Proximal Causal Learning with Kernels: Two-stage Estimation and Moment Restriction**

(***alphabetical order**) Afsaneh Mastouri*, **Yuchen Zhu***, Limor Gultchin, Anna Korba, Ricardo Silva, Matt Kusner, Arthur Gretton, Krikamol Muandet

Grants, Awards and Distinctions

2020-24   **EPSRC PhD Studentship (~£19000 + tuition fee / yr)**

2018   **SRIM Summer Research Grant, University of Cambridge (£2000)**

2016   **King's College Cambridge Travel Award.** For charity work with Education Partnerships Africa **(£1000)**

2016   **Bangor Rotary Club grant.** For charity work with Education Partnerships Africa **(£400)**

2015   **F.C.Baines Scholarship, David Hughes School (£300)**

2014   **British Mathematical Olympiad. (Silver Medal, ranked 36th in the UK.)**

2013, 2014   **UK Senior Maths Challenge 2013, 2014. (Gold, Gold)**

Service

2020 -   **Academic Service:** Reviewer and volunteer for machine learning conferences: ICML 2023-, UAI 2023-, NeurIPS 2022-, ICML SCIS workshop 2022, NeurIPS MLECON workshop 2021.

12.2015 - 09.2016   **Social Service:** Educational project management with Education Partnerships Africa.

Skills

Programming languages: Python (inc. PyTorch), Matlab (Advanced), R (basic). **Tools:** Git, LaTeX

Languages: English (Bilingual proficiency), Mandarin (Native)

Reference available upon request.