Yuen Chen

yuenc2@illinois.edu | chenyuen

Expected to graduate in Summer 2028. Looking for 2024 Research Scientist Intern or Research Intern Positions.

EDUCATION

University of Illinois at Urbana-Champaign — Aug 2023 - Expected May 2028
Ph.D. Student in Computer Science, advised by Prof. Hari Sundaram& Prof. Han Zhao
Research Topics: Causality, Bayesian Inference, Domain Generalization

University of California Berkeley — Aug 2020 - Dec 2022
B.A. Applied Mathematics & Statistics — GPA: 3.93/4.00

University of Copenhagen — Feb 2022 - Jun 2022
Exchange Student in Computer Science and Mathematical Science — GPA: 10.5/12

Irvine Valley College — Aug 2018 - May 2020
Associate in Science in Mathematics — GPA: 3.94/4.0

SKILLS

[Research] Causality, Machine Learning, Natural Language Processing

[Coding] Python, R/R Studio, MATLAB

[Others] PyTorch, TensorFlow, NumPy, Pandas, Prompt Engineering, Large-Language Model

RESEARCH EXPERIENCE

Graduate Researcher, University of Illinois, Urbana-Champaign — Aug 2023 - Present
Advisors: Prof. Hari Sundaram and Prof. Han Zhao
Conducted research on domain generalization to machine learning models' peformance under distribution shifts.
Accelerated stochastic variational inference through efficient sampling and tuning the mini-batch size.

Research Intern, Empirical Inference, Max Planck Institute for Intelligent Systems — Feb 2023 - Jul 2023
Supervisor: Prof. Bernhard Schölkopf & Zhijing Jin (Ph.D. at Max Planck Institute & ETH)
Worked in the intersection of causality and natural language processing, focusing on the causal reasoning ability of large-language models(LLM).
Developed prompting strategies that improved LLM's performance by 12 %.

Undergraduate Researcher, Language, Reasoning and Education Lab, ETH Zürich — Aug 2023 - Dec 2022
Supervisor: Prof. Mrinmaya Sachan & Zhijing Jin (Ph.D. at Max Planck Institute & ETH)
Conducted research on experimental design to optimize experimental tables with orthogonal arrays.

Undergraduate Researcher, University of Copenhagen — Feb 2022 - Jun 2022
Supervisor: Prof. Yevgeny Seldin
Conducted research to find an optimal algorithm for multi-armed bandits problems with time-based switching costs.

PAPERS

CLadder: Assessing Causal Reasoning in Language Models.
Zhijing Jin*, **Yuen Chen***, Felix Leeb*, Luigi Gresele*, Ojasv Kamal, Zhiheng Lyu, Kevin Blin, Fernando Gonzalez, Max Kleiman-Weiner, Mrinmaya Sachan, Bernhard Schölkopf. *arXiv 2023.*

Causal Impact Index: A Causal Formulation of Paper Impact.
Ishan Kumar Agrawal*, Zhijing Jin*, Ehsan Mokhtarian, Siyuan Guo, **Yuen Chen**, Negar Kiyavash, Mrinmaya Sachan, Bernhard Schölkopf. 2023.

HIGHLIGHTED PROJECTS

Time Series Analysis on Semiconductor Processing Tools — Fall 2022
Developed a machine learning model to predict the performance of wafer production process.
Condensed 560k+ data into 216 data by Fast Fourier Transform, sinusoidal regression, and quadratic regression.
Achieved 87% prediction accuracy on wafer metrology with a linear regression model.

Offline Evaluation of Bandit Algorithms — Spring 2022
Evaluated modified UCB1 and EXP3 with importance-weighted losses on "R6B Yahoo! Front Page Today Module User Click Log Datase".
Investigated the performance of UCB1, EXP3, and random strategy compared to the theoretical performance bound.
Achieved 2.6x better performance on EXP3 algorithm than the theoretical performance lower bound.

Medical Images Segmentation — Spring 2022
Implemented U-Net with PyTorch to segment the blood vessels on photographs of the retina.

*Equal Contribution

Trained the neural network with sample-splitting and output segmented images for the test image data.
Achieved <10% binary cross-entropy loss after 40 training epochs.

Representation Learning and Generative Modelling on MNIST Dataset — Spring 2022

Performed dimensionality reduction on MINIST dataset using PCA, Autoencoder (AE), and Variational Autoencoder (VAE).
Optimized binary cross-entropy loss of AE from 0.005 to 0.002 and evidence lower bound loss of VAE from 500+ to 200.
Synthesized new image data by sampling from the latent space of VAE.

R Packages Text Analysis — Spring 2021

Performed frequency text analysis by regular expression and visualized the outcome by ggplot2.
Built a web application with R-shiny that demos the analysis of a distribution of R packages' title length.

Modified Brent's Method — Sprint 2021

Implemented the root finding algorithm proposed by Wilkens and Gu in MATLAB.
Performed bisection method, inverse quadratic interpolation, and secant method to reduce the numbers of function evaluations.
· Achieved 92% test case coverage within tolerance of 1e -15 in terms of efficiency.

PROFESSIONAL & LEADERSHIP EXPERIENCE

Statistics Course Reader, University of California, Berkeley — Sep 2022 - Dec 2022

Assisted and graded 800+ students on homework, labs, exams weekly in Stat 20: Introduction to Statistics.
Co-managed 8 sections of lectures with Prof. Andrew Bray and other course staff on course content.

Math Tutor, Irvine Valley College — Aug 2019 - May 2022

Conducted 1-on-6 weekly sections and hosted office hours for 200+ students weekly.
Designed weekly problem sets for teaching uses in fields of trigonometry, calculus, and differential equations.
Earned 90% satisfaction in teaching from students in Fall 2021 and Spring 2022.

Undergraduate Researcher, Language, Reasoning and Education Lab, ETH Zürich — Aug 2023 - Dec 2022

Supervisor: Prof. Mrinmaya Sachan & Zhijing Jin (Ph.D. at Max Planck Institute & ETH)
Conducted research on experimental design to optimize experimental tables with orthogonal arrays.

Commissioner of Budget and Finance Committee, Student Government of Irvine Valley College — Aug 2018 - Sep 2019

Coordinated with 6 other committees on planning and advertising campus events.
Allocated $800,000 in funds to student organizations, intercurricular programs, and scholarships.