ZHIHENG LYU

cogito@connect.hku.hk ⋄ http://cogito233.github.io ⋄ (+86) 180 4225 1853

EDUCATION

The University of Hong Kong 2020 - 2024
B.Eng. in Computer Science *Hong Kong SAR, China*
- **1st Class Honours** (GPA: 3.76/4.3)
- Got Two Gold Medal in ICPC Regional (Nanjing & Shanghai), Entried 45th ICPC World Final

University of California, Berkeley 2022 - 2023
Undergraduate Special Student Program (GPA: 3.93/4.0) *Berkeley, CA, USA*
- Courses include **CS288** Natural Language Processing (A), **CS182** Deep Learning (A).

CURRENT POSITION

Undergraduate Researcher at Berkeley AI Research, *CA, USA* 2023 - present
Supervisor: Kevin Yang and Prof. Dan Klein

RESEARCH EXPERIENCE

Detecting Factual Contradiction in Structural Generated Story [Report] [Slides](Ongoing)
UCB BAIR, NLP Group, with **Prof. Dan Klein** & **Kevin Yang**
- Deconstructed complex narrative structures into **atomic fact statements** and designed an advanced data structure to model **dynamic world states**, with an mechanism for **detecting inner-plot factual contradictions** and **maintaining the time-bound fact validity** within the narrative framework.
- Refined the system through agile development and refactor; optimized hierarchical algorithms for scalability and precision; utilizing fine-tuned Hugging Face models for filtering; crafted a dedicated website for custom annotations with HTML, CSS, and JavaScript.

Probing the Understanding of Large Language Model in Causality
ETH Zürich, NLPED lab, with **Prof. Mrinmaya Sachan**, *Prof. Bernhard Schölkopf &* **Zhijing Jin**
- Analyzed LLM's understanding of causal relationships in the Sentiment Dataset [4] and its potential to improve predictions and discover causality by using causal direction in prompts [5];
- Design and Construct **corr2cause** dataset [1] to evaluate LLM's grasp of causality with developed C++ code with bitmask dynamic programming and graph isomorphism techniques.

Explore new Training Paradigm for a Causal Graph based on Conditional Independent Test
ETH Zürich, NLPED lab, with Prof. Mrinmaya Sachan, **Prof. Bernhard Schölkopf** *&* **Zhijing Jin**
- Delved into causality language, bias, and robustness, analyzed real-world and pseudo datasets for machine learning evaluations, and optimized training by implementing KCIT in PyTorch with gradient back propagation and test statistic normalization [Paper under progress].
- Experimented with causal graph-based neural network model and training (details on Here) and contributed to the open-source project CAUSAL-LEARN.

Construct and Pre-process Datasets on NLP for Social Good
University of Michigan, LIT, with **Prof. Rada Mihalcea** *&* **Zhijing Jin**
- Built a web crawler for AIScholar[3] and Logical Fallacy[6] using Selenium, with captcha bypass and exception handling. Organize the dataset, removing duplicates and errors.
- Developed the whole open-source project CAUSALNLP-TOOLBOX for data visualization, text matching, and causal experiment statistics.

Snapclf: An Automatic Object Classification Model Training System for AIoT Applications.
Megvii, R-face Institution, with Erjin Zhou
- Constructed a model searching pipeline on cluster with GPUs; Establish data pipelines and performance visualization using packages like **numpy**, **seaborn**, **pandas**, **sklearn** e.t.c; Accelerate python data preprocess speed by using **pybind11** and **opencv**; Wrap up and release python packages.

PUBLICATIONS

1. (ICLR 2024 Under Review) Can Large Language Models Infer Causation from Correlation?
   Zhijing Jin, Jiarui Liu, **Zhiheng Lyu***, Spencer Poff, Mrinmaya Sachan, Rada Mihalcea, Mona Diab, Bernhard Schölkopf.*
2. (NeurIPS 2023) CLADDER: Assessing Causal Reasoning in Language Models
   Zhijing Jin, Yuen Chen*, Felix Leeb*, Ojasv Kamal,* **Zhiheng Lyu***, Kevin Blin, Fernando Gonzalez, Max Kleiman-Weiner, Luigi Gresele, Mrinmaya Sachan, Bernhard Schölkopf.*
3. (ICWSM 2024 Under Review) Voices of Her: Analyzing Gender Differences in the AI Publication World
   Yiwen Ding, Jiarui Liu*,* **Zhiheng Lyu***, Kun Zhang, Bernhard Schoelkopf, Zhijing Jin, Rada Mihalcea.*
4. (NAACL 2024 In Preparation) Psychologically-Inspired Causal Prompts
   Zhiheng Lyu*, Zhijing Jin*, Rada Mihalcea, Mrinmaya Sachan, Bernhard Schölkopf.*

5. (UAI 2022 - Causal Representation Learning Workshop) Can Large Language Models Distinguish Cause from Effect?
Zhiheng Lyu*, Zhijing Jin*, Mrinmaya Sachan, Rada Mihalcea, Bernhard Schölkopf.
6. (EMNLP 2022) Logical Fallacy Detection
Zhijing Jin, Abhinav Lalwani*, Tejas Vaidhya, Xiaoyu Shen, Yiwen Ding, **Zhiheng Lyu**, Mrinmaya Sachan, Rada Mihalcea, Bernhard Schölkopf.*

INVITED TALK

Sept 22. 2023 "Challenges in ML System: Bias, Robustness and Adversarial Attack" Invited talk at HKU Innovation Wing organized by AI4Good for undergraduate engineering students

WORK EXPERIENCE

UC Berkeley, *Undergraduate Researcher*	Mar 2023 - now
ETH Zürich, *Undergraduate Researcher*	Sept 2021 - Jan 2023
R-face institution @ Megvii, *Machine Learning Engineer*	June 2021 - Aug 2021

TEACHING EXPERIENCE

HKU, *Student TA @ COMP3314 Machine Learning 2023Fall*	Sept 2023 - Dec 2023
HKU, *Trainer & Team Captain @ Algorithm Competition Group*	Sept 2020 - June 2021
Shaoxing No.1 High School, *Olympic Information Student Teacher*	July 2020 - Oct. 2020

LEADERSHIP EXPERIENCE

HKU AI4Good HKU, *Founder*	Sept 2023 - Now
AStar Team @ HKU, *Team Leader*	July 2021 - Aug 2022

AWARDS & HONORS

1. Ranked 50 in **45th International Collegiate Programming Contest (ICPC) World Finals**, Dhaka, Nov. 2022
2. 9th Place Team in **ICPC World Challenge**, Dhaka, Nov. 2022
3. Gold Medalist and Team Captain in ICPC 2021 Shanghai region – ranked 20/668,Shanghai,Dec. 2021
4. Gold Medalist and Team Captain in ICPC 2020 Nanjing region – ranked 7/556, Nanjing, Dec. 2020
5. Silver Medalist at National Olympiad in Informatics (NOI) 2018, Changsha, Jul. 2018
6. Gold Medalist at The Asia-Pacific Informatics Olympiad (APIO) 2018, Beijing, Apr. 2018

GRANTS & SCHOLARSHIP

Grants
- 150,000 HKD, Funding Scheme for student projects/activities by Tam Wing Fan Innovation Fund and Philomathia Foundation Innovation Fund, AStar @ HKU
- 50,000 HKD, Gallant Ho Experiential Learning fund, AStar @ HKU

Scholarships
- 40,000 HKD, Undergraduate Research Fellowship Programme (URFP) @ HKU
- 24,000 HKD, The Lee Shau Kee Scholarships for Student Enrichment @ HKU

PROFESSIONAL SERVICE

Co-organizer, a half-day workshop *Performance Optimization in the LLM world* @ ICPE 2024.
Reviewer, the 2024 Annual Conference of the Association for Computational Linguistics. ACL 2024.

ADVISING & MENTORING

- Yucheng Lin, 2023-now. Undergraduate in FinTech at The University of Hong Kong.

REFERENCES

Kevin Yang, PhD Candidate, Berkeley AI Research, UC Berkeley & Meta AI
Jin Zhijing, PhD Candidate, Max Planck Institute & ETH, affiliated at UMich
Prof. Kong Lingpeng, Assistant Professor, Dept. of Computer Science, University of Hong Kong
Prof. Mrinmaya Sachan, Assistant Professor, Dept. of Computer Science, ETH Zürich
Prof. Dan Klein, Professor, Dept. of Computer Science, UC Berkeley