

Min-Hsuan (Samuel) Yeh

myeh4@wisc.edu

mhyeh.github.io

+1 (413) 768-7441

Research Interest

Reliable machine learning and Natural language processing

Education

- 2024– *Ph.D. in Computer Science*
University of Wisconsin-Madison, Madison, WI
• Advisor: [Prof. Sharon Yixuan Li](#)
- 2022–24 *M.S. in Computer Science*
University of Massachusetts-Amherst, Amherst, MA
- 2016–20 *B.Sc. in Computer Science*
National Taiwan University of Science and Technology, Taipei, Taiwan

Research Experience

- 2023–24 *Research Assistant Intern*
[Crowd-AI Lab](#), PSU, University Park, PA, USA (Advisor: [Prof. Ting-Hao Huang](#))
• Research on *logical fallacy detection*
- 2022–23 *Independent Study*
UMass-Amherst, Amherst, MA, USA (Advisor: [Prof. Philip S. Thomas](#))
• Research on *fair machine learning*
- 2020–22 *Research Assistant*
[NLPSA Lab](#), Academia Sinica, Taipei, Taiwan (Advisor: [Dr. Lun-Wei Ku](#))
• Research on *visual question generation and lie detection*
- 2017–20 *Undergraduate Research Assistant*
[GAME Lab](#), NTUST, Taipei, Taiwan (Advisor: [Dr. Wen-Kai Tai](#))
• Research on *game design and reinforcement learning*

Publications

 [Google Scholar](#)

Peer-reviewed Conference Proceedings

5. **Yeh, Min-Hsuan**, Metevier, B., Hoag, A. & Thomas, P. *Analyzing the Relationship Between Difference and Ratio-Based Fairness Metrics* in *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (2024).
4. **Yeh, Min-Hsuan**, Wan, R. & Huang, T.-H. *CoCoLoFa: A Dataset of News Comments with Common Logical Fallacies Written by LLM-Assisted Crowds* in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (2024).

3. **Yeh, Min-Hsuan**, Chen, V., Huang, T.-H. & Ku, L.-W. *Multi-VQG: Generating Engaging Questions for Multiple Images in Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* (2022).
2. **Yeh, Min-Hsuan** & Ku, L.-W. *Lying Through One's Teeth: A Study on Verbal Leakage Cues in Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (2021).

Working papers

1. **Yeh, Min-Hsuan.**, Tao, L., Wang, J., Du, X. & Li, Y. *How Reliable Is Human Feedback For Aligning Large Language Models?* Under Review. 2024.

Academic Service

Reviewer

2025 COLING
2024 ACL, COLING

Last updated: October 3, 2024