Han Liu

https://hanliuai.github.io liuhantech@gmail.com (314) 349-4914

Expertise

I develop and deploy scalable machine learning systems for human-AI collaboration. My work includes fine-tuning large language models, optimizing deep learning for vision and multimodal learning, and ensuring interpretability, robustness, and efficiency for real-world applications.

Education

University of Chicago September 2020 - August 2024
Ph.D. in Computer Science

Washington University in St. Louis August 2015 - May 2019
B.A. in Mathematics, Computer Science, and Minor in Linguistics

Experience

University of Chicago September 2020 - Present
Research Assistant (@ Chicago Human+AI Lab) *Chicago, IL*

Project 1: AI-driven Tutorials for Image Classification

- Developed deep learning models integrating task supervision and human perception signals, improving non-expert accuracy by 13-17% with model explanations in medical and natural image classification. [3]
- Finetuned vision language models (VLMs) for image understanding tasks, deriving text-based explanations from VLMs to assist human decision-making.

Project 2: Prostate Cancer Diagnosis with AI

- Developed models for detecting prostate cancer from MRI scans, securing a top-6 placement in the PICAI challenge, and interfaces for human-AI collaboration demoed at RSNA 2023 Education Exhibit.
- Conducted a cross-institutional study with radiologists, revealing under-reliance on AI and opportunities to improve human-AI collaboration. [1]

Project 3: Improving Value Alignment of Large Language Models

- Investigated limitations of previous reinforcement learning from human preferences (RLHF) algorithms under conditions of heterogeneous human preference data.
- Extended the direct preference optimization (DPO) algorithm with diverse divergence constraints. [2]

Microsoft Research June 2022 - September 2022
Research Intern (@ Human-AI eXperiences (HAX) Team) *Redmond, WA*

- Studied how offline metrics of code generation models align with human judgments, proposing a hybrid metric that better reflects programmer-perceived value. [4]

University of Colorado Boulder August 2019 - August 2020
Research Assistant (@ NLP+CSS Lab) *Boulder, CO*

- Studied how human and AI collaborate and complement each other under the effect of distribution shift and interactive interfaces in various decision making tasks such as deceptive review detection, profession classification, and recidivism prediction. [5]
- Analyzed results from large-scale human experiments to study how different types of model-driven tutorials and real-time assistance from model explanations help humans in decision making tasks. [6]

Selected Publications

[1] Chacha Chen, **Han Liu**, Jiamin Yang, Benjamin M. Mervak, Bora Kalaycioglu, Grace Lee, Emre Cakmakli, Matteo Bonatti, Sridhar Pudu, Osman Kahraman, Gül Gizem Pamuk, Aytekin Oto, Aritrick Chatterjee, and Chenhao Tan. Can Domain Experts Rely on AI Appropriately? A Case Study on AI-Assisted Prostate Cancer MRI Diagnosis, Preprint (2025).

[2] Chaoqi Wang, Yibo Jiang, Chenghao Yan, **Han Liu**, and Yuxin Chen. Beyond Reverse KL: Generalizing Direct Preference Optimization with Diverse Divergence Constraints. In *International Conference on Learning Representations*, (Spotlight, ICLR 2024).

[3] **Han Liu**, Yizhou Tian, Chacha Chen, Shi Feng, Yuxin Chen, and Chenhao Tan. Learning Human-Compatible Representations for Case-Based Decision Support. In *International Conference on Learning Representations*, (ICLR 2023).

[4] Victor Dibia, Adam Fourney, Gagan Bansal, Forough Poursabzi-Sangdeh, **Han Liu**, and Saleema Amershi. Aligning Offline Metrics and Human Judgments of Value of AI-Pair Programmers. In *Findings of the Association for Computational Linguistics: ACL 2023*, (Findings of ACL 2023).

[5] **Han Liu**, Vivian Lai, and Chenhao Tan. Understanding the Effect of Out-of-distribution Examples and Interactive Explanations on Human-AI Decision Making. *Proceedings of the ACM on Human-Computer Interaction, Volume 5, Issue CSCW2*, (CSCW 2021).

[6] Vivian Lai, **Han Liu**, and Chenhao Tan. "Why is 'Chicago' deceptive?" Towards Building Model-Driven Tutorials for Humans. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, (CHI 2020).

More publications can be found on Google Scholar.

Technical Skills

Programming Languages	Python, Java, C++, SQL, JavaScript, HTML, CSS
Machine Learning Tools	PyTorch, Keras, Scikit-learn, Pandas, NumPy, Jupyter
Web Development	React, Node.js, Vue.js, jQuery, Django, Bootstrap, MySQL, SQLite

Experience with training and inference with Large Language Models (LLMs) and Vision-Language Models (VLMs): LLaMA-2, Vicuna, LLaVA, CLIP etc.

Teaching Experience

Teaching assistant for the following courses (selected):

CMSC 25100: Introduction to Machine Learning (UChicago, Winter 2024)
CMSC 35400: Machine Learning (UChicago, Sprint 2024)
CSCI 5622: Machine Learning (CU Boulder, Fall 2019)
CSE 559A: Computer Vision (WUSTL, Fall 2018)
CSE 511A: Introduction to Artificial Intelligence (WUSTL, Fall 2018)