

# Yifei He | Curriculum Vitae

✉ yifeihe3@illinois.edu • 🌐 yifei-he.github.io

My research interest is centered around trustworthy machine learning. I currently work on i) improving the multi-task capabilities of foundation models and ii) multi-objective LLM alignment. Previously, I have also worked on multimodal learning and domain adaptation/generalization.

## Education

### University of Illinois Urbana-Champaign (UIUC)

*Ph.D. in Computer Science*

*M.S. in Computer Science*

Advisor: Prof. Han Zhao

**Urbana, IL, USA**

*May 2023 - Present*

*Aug 2021 - May 2023*

### University of Michigan (UM)

*B.S.E. in Data Science, minor in Mathematics*

Summa Cum Laude

**Ann Arbor, MI, USA**

*Aug 2019 - Apr 2021*

### Shanghai Jiao Tong University (SJTU)

*B.S.E. in Electrical and Computer Engineering*

**Shanghai, China**

*Sept 2017 - Aug 2021*

## Industry Experience

### Microsoft

*Research Intern, GenAI*

Manager: Dr. Yang Liu

- Worked on distillation of Mixture-of-Experts (MoE) models.

**Redmond, WA, USA**

*Aug 2024 - present*

### Microsoft

*Applied Scientist Intern, Turing*

Manager: Dr. Alon Benhaim

- Worked on scaling law for multilingual language models.

**Redmond, WA, USA**

*May 2024 - Aug 2024*

### Amazon

*Applied Scientist Intern, Search Science and AI*

Manager: Dr. Alejandro Mottini

- Improved large-scale multi-task tuning of foundation models.
- Developed a vision-language retrieval foundation model with instruction tuning.

**Seattle, WA, USA**

*May 2023 - Aug 2023*

## Publications (\* denotes equal contribution)

- [1] **Localize-and-Stitch: Efficient Model Merging via Sparse Task Arithmetic.**  
Yifei He, Yuzheng Hu, Yong Lin, Tong Zhang, Han Zhao.  
*Under review.*
- [2] **Semi-Supervised Reward Modeling via Iterative Self-Training.**  
Yifei He\*, Haoxiang Wang\*, Ziyang Jiang, Alexandros Papangelis, Han Zhao  
In *Findings of the Association for Computational Linguistics: EMNLP 2024*. (**EMNLP 2024 Findings**)
- [3] **Robust Multi-Task Learning with Excess Risks.**  
Yifei He, Shiji Zhou, Guojun Zhang, Hyokun Yun, Yi Xu, Belinda Zeng, Trishul Chilimbi, Han Zhao.  
In *Proceeding of the 41st International Conference on Machine Learning*. (**ICML 2024**)
- [4] **Gradual Domain Adaptation: Theory and Algorithms.**  
Yifei He\*, Haoxiang Wang\*, Bo Li, Han Zhao.  
In *Journal of Machine Learning Research*. (**JMLR 2024**)

- [5] **Efficient Modality Selection in Multimodal Learning.**  
**Yifei He\***, Runxiang Cheng\*, Gargi Balasubramaniam\*, Yao-Hung Hubert Tsai, Han Zhao.  
 In *Journal of Machine Learning Research*. (**JMLR 2024**)  
 (Extended version of publication [6].)
- [6] **Greedy Modality Selection via Approximate Submodular Maximization.**  
 Runxiang Cheng\*, Gargi Balasubramaniam\*, **Yifei He\***, Yao-Hung Hubert Tsai, Han Zhao.  
 In *Proceedings of the 38th Conference on Uncertainty in Artificial Intelligence*. (**UAI 2022**)
- [7] **Conformer-RL: A Deep Reinforcement Learning Library for Conformer Generation.**  
 Runxuan Jiang, Tarun Gogineni, Joshua Kammeraad, **Yifei He**, Ambuj Tewari, Paul Zimmerman.  
 In *Journal of Computational Chemistry*. (**JCC 2022**)
- [8] **A Hierarchical Approach to Multi-Event Survival Analysis.**  
 Donna Tjandra, **Yifei He**, Jenna Wiens.  
 In *Proceedings of the 35th AAAI Conference on Artificial Intelligence*. (**AAAI 2021**)

## Professional Service

---

**Reviewer:** UAI, NeurIPS, ICLR, AISTATS, ICML

## Teaching Experience

---

Teaching assistant at UIUC

- CS 357 Numerical Methods I 2022 Fall, 2022 Spring
- CS 441 Applied Machine Learning 2021 Fall

Teaching assistant at UM

- EECS 445 Intro to Machine Learning 2020 Fall

## Skills

---

**Programming:** Python, Java, C++, Matlab, R,  $\text{\LaTeX}$ , Mathematica

**Framework:** PyTorch, DeepSpeed, TensorFlow, Keras, Gym