Yifei He | Curriculum Vitae

(217) 377 3761 • yifeihe3@illinois.edu • yifei-he.github.io

My research goal is to build efficient and robust machine learning systems. I work on trustworthy machine learning, including multi-task learning, multimodal learning and domain adaptation/generalization.

Education

University of Illinois Urbana-Champaign (UIUC) **Urbana, IL, USA**
Ph.D. in Computer Science *May 2023 - Present*
M.S. in Computer Science *Aug 2021 - May 2023*
Advisor: Prof. Han Zhao

University of Michigan (UM) **Ann Arbor, MI, USA**
B.S.E. in Data Science, minor in Mathematics *Aug 2019 - Apr 2021*
Summa Cum Laude

Shanghai Jiao Tong University (SJTU) **Shanghai, China**
B.S.E. in Electrical and Computer Engineering *Sept 2017 - Aug 2021*

Industry Experience

Amazon **Seattle, WA, USA**
Applied Scientist Intern, Search Science and AI *May 2023 - Present*

- Developed a vision-language retrieval foundation model with instruction tuning.

Publications

(* denotes equal contribution)

[1] Robust Multi-Task Learning with Excess Risks.
Yifei He, Shiji Zhou, Guojun Zhang, Hyokun Yun, Yi Xu, Belinda Zeng, Trishul Chilimbi, Han Zhao
Under review.

[2] Generative Gradual Domain Adaptation with Optimal Transport.
Yifei He*, Haoxiang Wang*, Han Zhao.
In *ICML Principles of Distribution Shift (PODS) Workshop, 2022*.

[3] Efficient Modality Selection in Multimodal Learning.
Yifei He*, Runxiang Cheng*, Gargi Balasubramaniam*, Yao-Hung Hubert Tsai, Han Zhao.
Under review.

[4] 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)

[5] 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)

[6] 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*.

Research Experience

Multi-Objective Optimization for Robust Multi-task Learning **Urbana, IL, USA**
Advisor: Prof. Han Zhao, Department of Computer Science, UIUC *July 2022 - Present*

- Developed an adaptive task balancing algorithm using excess risk estimation to address task noise.
- First-authored paper under review.

Generative Gradual Domain Adaptation **Urbana, IL, USA**
Advisor: Prof. Han Zhao, Department of Computer Science, UIUC *Mar 2022 - Sept 2022*

- Developed a framework to generate intermediate domains, improving domain adaptation under large distribution shift, alleviating the burden of data collection and extending the applications of gradual domain adaptation (GDA).
- First-authored publication at ICML PODS workshop 2022.

Modality Selection in Multimodal Learning **Urbana, IL, USA**
Advisor: Prof. Han Zhao, Department of Computer Science, UIUC *Aug 2021 - June 2022*

- Theoretically proved how to select the most informative subset of modalities given computational constraints.
- First-authored publication at UAI 2022.

Reinforcement Learning (RL) for Sequential Conformer Search **Ann Arbor, MI, USA**
Advisor: Prof. Ambuj Tewari, Department of Statistics, UM *July 2020 - Apr 2021*

- Applied RL algorithms to efficiently find the most stable structure of large molecules.
- Publication at Journal of Computational Chemistry.
- Open-sourced Python library "conformer-rl".

Deep Learning for Multi-Event Survival Analysis **Ann Arbor, MI, USA**
Advisor: Prof. Jenna Wiens, Department of Computer Science, UM *Apr 2020 - Sept 2020*

- Applied multi-task and hierarchical learning to better model the inter-event relations in survival analysis.
- Publication at AAAI 2021.

Professional Service

Reviewer: UAI 2023, NeurIPS 2023

Teaching Experience

CS 357 Numerical Methods I (UIUC) 2022 Fall, 2022 Spring
CS 441 Applied Machine Learning (UIUC) 2021 Fall
EECS 445 Intro to Machine Learning (UM) 2020 Fall

Skills

Specialties: Transfer Learning, Multimodal Learning, Foundation Models
Programming: Python, Java, C++, Matlab, R, SQL, JavaScript, HTML, LaTeX, Mathematica
Framework: PyTorch, TensorFlow, Keras, AWS, Gym, Hadoop