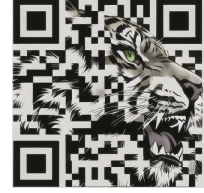


## JIAXUAN WANG

Tel: (734) 834-7996

E-mail: [jiaxuan@umich.edu](mailto:jiaxuan@umich.edu)

Website: [nathanwang000.github.io](https://nathanwang000.github.io)



### EDUCATION

---

#### Ph.D. at the University of Michigan, Ann Arbor

Sep 2017-April 2022

*Computer Science and Engineering*

Advisor: [Jenna Wiens](#)

GPA: 4.00 / 4.00

Research interests: Model interpretability; Alignment; Time-series analysis; Transfer/multitask learning; Non convex optimization; Temporal conditional shift; Computer vision; Deep reinforcement learning; Causal inference; Basketball analytics

Computational skills: PyTorch; Python; C++; Javascript; Matlab; R;

#### Bachelors of Science in Engineering, Ann Arbor

Sep. 2013 - Dec. 2016

*Computer Science major and Mathematics minor*

GPA: 3.96 / 4.00

Directed research: Computer vision; Basketball analytics

### EMPLOYMENT

---

#### Senior AI research scientist, GE Healthcare

April 8 2024 - Present

Developing AI algorithms to improve perinatal health outcome.

#### Research scientist in machine learning, Meta

Jun. 6 2022 - April 1 2024

- With teammates, developed a reinforcement learning agent to protect user data by selecting appropriate actions on scrappers (e.g., blocking, reCAPTCHA or SMS challenges). This is the first known RL system applied to the user abuse problem, significantly outperforming the traditional classify then enforce paradigm in reducing data scraped without degrading user experience [paper in submission].
- Proposed and implemented a feature attribution framework to debug and monitor distribution shift for the reinforcement learning agent used in production.
- With teammates, developed the duo-llama system that utilized a fine-tuned large language model (LLM) to save human labeling cost while enhancing LLM's ability on low resource languages through machine translation. The resulting system significantly outperforms the production classifier on 4 already supported languages, while enabling the new classifier to handle 17 languages.

#### Research Intern, Adaptive Systems and Interaction Group, Microsoft Research

Jun. 1 - Aug.21 2020

Mentor: [Scott Lundberg](#)

Proposed a novel explanation method, Shapley Flow, that unifies and avoids the pitfall of 3 previous methods.

#### Software Engineering Intern, NLP group, Bloomberg L.P. (New York)

Jun. 7 - Aug.19 2016

Mentors: [Konstantine Arkoudas](#) and [Srivas Prasad](#)

Algorithms for natural language parsing in financial chart domain: C++; SVM; PCFG

#### Research Assistant, Computer vision lab, University of Michigan

Oct. 2014 - Jan. 2016

Advisor: [Jia Deng](#)

Focus: Human action dataset collection; Amazon Mechanical Turk; Feature extraction; Rotation equivariant network

## PUBLICATIONS (\* denotes equal contribution)

---

### 1. [Learning Concept Credible Models for Mitigating Shortcuts](#)

**Jiaxuan Wang**, Sarah Jabbour, Maggie Makar, Jenna Wiens

Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS), 2022

### 2. [Shapley Flow: A Graph-based Approach to Interpreting Model Predictions](#)

**Jiaxuan Wang**, Jenna Wiens, Scott Lundberg

Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS), 2021

### 3. [AdaSGD: Bridging the gap between SGD and Adam](#)

**Jiaxuan Wang**, Jenna Wiens

arXiv preprint, 2020

### 4. [Relaxed Parameter Sharing: Effectively Modeling Time-Varying Relationships in Clinical Time-Series](#)

Jeeheh Oh\*, **Jiaxuan Wang\***, Shengpu Tang, Michael Sjoding, Jenna Wiens

In Proceedings of the 4th Machine Learning for Healthcare Conference, 2019

### 5. [Learning Credible Models](#)

**Jiaxuan Wang**, Jeeheh Oh, Haozhu Wang, Jenna Wiens

ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2018

### 6. [The Advantage of Doubling: A Deep Reinforcement Learning Approach to Studying the Double Team](#)

**Jiaxuan Wang\***, Ian Fox\*, Jonathan Skaza, Nick Linck, Satinder Singh, Jenna Wiens

MIT Sloan Sports Analytics Conference, 2018

### 7. [Learning to Exploit Invariances in Clinical Time-Series Data using Sequence Transformer Networks](#)

Jeeheh Oh, **Jiaxuan Wang**, and Jenna Wiens

In Proceedings of the 4th Machine Learning for Healthcare Conference, 2018

### 8. [HICO: A Benchmark for Recognizing Human-Object Interactions in Images](#)

Yu-Wei Chao, Zhan Wang, Yugeng He, **Jiaxuan Wang**, Jia Deng

International Conference on Computer Vision (ICCV) 2015

## TECHNICAL REPORTS

---

### 1. [Using feature attribution to debug and monitor distribution shift for a production ML system](#)

**Jiaxuan Wang**

2023

## SERVICES

---

Reviewer @ MLHC 2023

Reviewer @ NeurIPS 2022

Reviewer @ AISTATS 2022

Reviewer @ ICLR 2022

Reviewer @ AISTATS 2021

Reviewer @ NeurIPS 2020

Reviewer @ NeurIPS 2019

Reviewer @ MLHC 2021

Reviewer @ MLHC 2020  
Volunteer @ Michigan AI symposium 2020  
Reviewer @ SSAC 2020  
Reviewer @ MLHC 2019  
Reviewer @ SSAC 2019  
Volunteer @ Michigan AI symposium 2019