RESEARCH INTERESTS
Sustainable Computing, AI/ML Systems, Datacenter Computing, Mental Health

PROFESSIONAL EXPERIENCE

Purdue University
Assistant Professor in Elmore Family School of Electrical and Computer Engineering
PI, STYLE (Sustainable computing Systems and LEarning) Lab
West Lafayette, IN, USA
8/2023 – Present

Massachusetts Institute of Technology
Postdoctoral Associate & NSF Computing Innovation Fellow. Mentor: Michael Carbin
Cambridge, MA, USA
1/2021 – 8/2023

Meta
Visiting Researcher
Cambridge, MA, USA
10/2021–12/2022

Google
Research Intern
Sunnyvale, CA, USA
6/2019-9/2019

EDUCATION

University of Chicago
Ph.D. & MS in Computer Science. Advisor: Henry Hoffmann
Chicago, IL, USA
8/2015 – 12/2020

Nanyang Technological University
Ph.D. Candidate in Computer Science. Passed Qualification Exam.
Singapore
7/2013 – 7/2015

Beijing Jiaotong University
B.E. in Electronic Science and Technology. Graduated with Highest Honor.
Beijing, China
9/2008 – 6/2012

SELECTED AWARDS AND HONORS

Innovation Award, Quantum Computing for Drug Discovery Challenge at ICCAD 2023
CRA/CCC/NSF Computing Innovation Fellowship 2020-2023
Meta Research Award 2021
EECS Rising Stars at UC Berkeley 2020

PUBLICATIONS AND PRESENTATIONS

Google Scholar
† Students mentored by me; ⋆ Equal contribution; ‡ Corresponding faculty author

Peer-reviewed Conference Proceedings

[C1] Yi Ding† and Tianyao Shi†, “Sustainable LLM Serving: Environmental Implications, Challenges, and Opportunities”. In: Proceedings of the 15th International Green and Sustainable Computing Conference (IGSC). 2024.

[C2] Amy Li†, Sihang Liu, and Yi Ding†, “Uncertainty-Aware Decarbonization for Datacenters”. In: Proceedings of the 3rd Workshop on Sustainable Computer Systems (HotCarbon). 2024.

[C3] Sophia Nguyen⋆,†, Beihao Zhou⋆,†, Yi Ding, and Sihang Liu. “Towards Sustainable Large Language Model Serving”. In: Proceedings of the 3rd Workshop on Sustainable Computer Systems (HotCarbon). 2024.

[C4] Gokul Subramanian Ravi, Pranav Gokhale, Yi Ding, William Kirby, Kaitlin Smith, Jonathan M Baker, Peter J Love, Henry Hoffmann, Kenneth R Brown, and Frederic T Chong. “CAFQA: A classical simulation bootstrap for variational quantum algorithms”. In: Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS). 2023.

- 2023 Innovation Award, Quantum Computing for Drug Discovery Challenge at ICCAD.
[C5] Alex Renda, Yi Ding, and Michael Carbin. “Turaco: Complexity-Guided Data Sampling for Training Neural Surrogates of Programs”. In: Proceedings of the ACM on Programming Languages (OOPSLA). 2023.

[C6] Yi Ding, Avinash Rao, Hyebin Song, Rebecca Willett, and Henry Hoffmann. “NURD: Negative-Unlabeled Learning for Online Datacenter Straggler Prediction”. In: Proceedings of Machine Learning and Systems (MLSys). 2022.

[C7] Yi Ding, Ahsan Pervaiz, Michael Carbin, and Henry Hoffmann. “Generalizable and interpretable learning for configuration extrapolation”. In: Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE). 2021.

[C8] Alex Renda, Yi Ding, and Michael Carbin. “Programming with neural surrogates of programs”. In: Proceedings of the 2021 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software (Onward!) 2021.

[C9] Yi Ding and Panos Toulis. “Dynamical systems theory for causal inference with application to synthetic control methods”. In: International Conference on Artificial Intelligence and Statistics (AISTATS). 2020.

[C10] Ming Gao, Yi Ding, and Bryon Aragam. “A polynomial-time algorithm for learning nonparametric causal graphs”. In: Advances in Neural Information Processing Systems (NeurIPS). 2020.

[C11] Yi Ding, Nikita Mishra, and Henry Hoffmann. “Generative and multi-phase learning for computer systems optimization”. In: Proceedings of the 46th International Symposium on Computer Architecture (ISCA). 2019.

[C12] Yi Ding, Risi Kondor, and Jonathan Eskreis-Winkler. “Multiresolution kernel approximation for Gaussian process regression”. In: Advances in Neural Information Processing Systems (NeurIPS). 2017.

[C13] Yi Ding, Chenghao Liu, Peilin Zhao, and Steven CH Hoi. “Large scale kernel methods for online auc maximization”. In: 2017 IEEE International Conference on Data Mining (ICDM). 2017.

[C14] Yi Ding, Peilin Zhao, Steven Hoi, and Yew-Soon Ong. “An adaptive gradient method for online auc maximization”. In: Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). 2015.

[C15] Pengcheng Wu, Yi Ding, Peilin Zhao, Chunyan Miao, and Steven Hoi. “Learning relative similarity by stochastic dual coordinate ascent”. In: Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). 2014.

Peer-reviewed Journals

[J1] Guillaume Basse, Yi Ding, and Panos Toulis. “Minimax designs for causal effects in temporal experiments with treatment habituation”. In: Biometrika. 2023.

[J2] Kathryn E Schertz, James Saxon, Carlos Cardenas-Iniguez, Luís Bettencourt, Yi Ding, Henry Hoffmann, and Marc G Berman. “Neighborhood street activity and greenspace usage uniquely contribute to predicting crime”. In: Npj Urban Sustainability. 2021.

Workshop Presentations

[W1] Yi Ding, Avinash Rao, and Henry Hoffmann. “Causal and Interpretable Learning for Datacenter Latency Prediction”. In: Women in Machine Learning Workshop co-located with NeurIPS (WiML) (2020).

[W2] Ming Gao, Yi Ding, and Bryon Aragam. “A Polynomial-time Algorithm for Learning Nonparametric Causal Graphs”. In: Women in Machine Learning Workshop co-located with NeurIPS (WiML) (2020).

[W3] Guillaume Basse, Yi Ding, and Panos Toulis. “Minimax Crossover Designs for Digital Experimentation”. In: Conference on Digital Experimentation at MIT (CODE@MIT) (2019).

[W4] Yi Ding, Guillaume Basse, and Panos Toulis. “Minimax Crossover Designs”. In: NeurIPS Workshop on “Do the right thing”: machine learning and causal inference for improved decision making (CausalML) (2019).
RESEARCH ADVISING

PhD Students
Tianyao Shi, Purdue University  
Yanran Wu, Purdue University  
William Meng, University of Pennsylvania

Master Students
Ashutosh Sharma, UIUC  
Hyunji Kim, MIT

Undergraduate Students
Sarah Deniz, Purdue University (DUIRI)  
Gavin Fortwendel, Purdue University (DUIRI)  
Yutao Han, University of Waterloo  
Leyi Yan, University of Waterloo  
Linda Wang, University of Waterloo  
Zihan Pan, University of Waterloo  
Amy Li, University of Waterloo (One HotCarbon’24 paper published)  
Beikiao Zhou, University of Waterloo (One HotCarbon’24 paper published)  
Sophia Nguyen, University of Waterloo (One HotCarbon’24 paper published)  
Avinash Rao, University of Chicago (One MLSys’22 paper published)

GRANTS

Title: Conference: DESC: Type III: A Holistic AI Computing Framework: Incorporating the Water and Biodiversity Dimensions of Sustainability  
Funder: NSF  
Duration: 2024–2025  
People: Inez Hua (PI), Yi Ding (Co-PI)  
Awarded: $9,9992 (My share: 50%)  

Title: Computing Innovation Fellows 2020 Project  
Funder: NSF  
Duration: 2020–2023  
People: Michael Carbin (PI), Yi Ding  
Awarded: $295,704  

Title: Meta Research Award on Statistics for Improving Insights, Models, & Decisions  
Funder: Meta  
Duration: 2021–2022  
People: Michael Carbin (PI), Yi Ding (Co-PI)  
Awarded: $46,000

TEACHING

Instructor, Purdue University, West Lafayette, IN  
Machine Learning in Cloud Computing (ECE 69500)  
Python for Data Science (ECE 20875)  
Python for Data Science (ECE 20875)

Teaching Assistant, University of Chicago, Chicago, IL  
Machine Learning and Large Scale Data Analysis (CMSC 25025)
PROFESSIONAL SERVICE

Organizer
NSF Workshop on Sustainable Computing: AI, Water, and Biodiversity, Co-Chair 2024

Program Committee
IEEE/ACM International Symposium on Computer Architecture (ISCA) 2025
IEEE International Symposium on High-Performance Computer Architecture (HPCA) 2025
USENIX Annual Technical Conference (ATC) 2024
Conference on Systems and Machine Learning (MLSys) 2024
ACM Student Research Competition at PACT 2023
SPLASH Onward! 2022
Conference on Systems and Machine Learning (MLSys) 2022
ACM Asia-Pacific Workshop on Systems 2022
Journal of Systems Research 2022

Technical Reviewing
Neural Information Processing Systems (NeurIPS) 2022
International Conference on Learning Representations (ICLR) 2022
International Conference on Machine Learning (ICML) 2022
Neural Information Processing Systems (NeurIPS) 2021
AAAI Conference on Artificial Intelligence (AAAI) 2021
AAAI Conference on Artificial Intelligence (AAAI) 2020
Neural Information Processing Systems (NeurIPS) 2019
International Conference on Machine Learning (ICML) 2019

PRESENTATIONS

Invited Seminars
Towards Sustainable Next Generation AI and Cloud Systems
Meta, Sunnyvale, USA Sep. 2024

A Holistic View on Machine Learning for Systems
University of Waterloo, Department of Computer Science Jun. 2023
Microsoft Research Apr. 2023
Texas A&M University, Department of Computer Science & Engineering Apr. 2023
University of Southern California, Department of Electrical & Computer Engineering Apr. 2023
University of Illinois, Department of Computer Science Mar. 2023
Cornell Tech, Department of Electrical & Computer Engineering Mar. 2023
Washington University in St. Louis, Department of Computer Science & Engineering Mar. 2023
Purdue University, School of Electrical & Computer Engineering Mar. 2023
Purdue University, Department of Computer Science Mar. 2023
Virginia Tech, Department of Computer Science Mar. 2023
Indiana University Bloomington, Department of Computer Science Feb. 2023
University of Colorado Boulder, Department of Computer Science Feb. 2023
University of Massachusetts Amherst, College of Information and Computer Sciences Feb. 2023

Conference Presentations
Uncertainty-Aware Decarbonization for Datacenters
Conference presentation at HotCarbon, Santa Cruz, USA Jul. 2024

Uncertainty-Aware Carbon Optimization in Cloud Computing
Conference presentation at SoDec Workshop at E-Energy, Singapore Jun. 2024
NURD: Negative-Unlabeled Learning for Online Datacenter Straggler Prediction
Conference presentation at MLSys, Santa Clara, USA Aug. 2022

Predictable Maintenance Job Planning in Datacenters
Meta Infrastructure Data Science Faculty Workshop at KDD, DC, USA Aug. 2022

Generalizable and Interpretable Learning for Configuration Extrapolation
Conference presentation at ESEC/FSE, Virtual Nov. 2021

Dynamical Systems Theory for Causal Inference with Application to Synthetic Controls
Causal Data Science Meeting, Virtual Nov. 2020
Conference presentation at AISTATS, Virtual Aug. 2020

Generative and Multi–phase Learning for Computer Systems Optimization
Conference presentation at ISCA, Phoenix, USA Jun. 2019

Multiresolution Kernel Approximation for Gaussian Process Regression
Conference presentation at NeurIPS, Long Beach, USA Dec. 2017

Large Scale Kernel Methods for Online AUC Maximization
Conference presentation at ICDM, New Orleans, USA Nov. 2017

An Adaptive Gradient Method for Online AUC Maximization
Conference presentation at AAAI, Austin, USA Jan. 2015

Last updated October 27, 2024