Miruna Oprescu

PhD Candidate in Computer Science
Cornell University, Cornell Tech
miruna@cs.cornell.edu | mirunaoprescu.com

INTERESTS Causal inference, machine learning, robust data-driven decision-making.

EDUCATION

Cornell University, Cornell Tech Fall 2021 – Present
Ph.D. Candidate in Computer Science. GPA: 4.00
Department of Energy Computational Science Graduate Fellow
M.S. in Computer Science, August 2024 (awarded en route to Ph.D.)

Harvard University May 2015
Joint A.B. degree in Physics and Mathematics. Minor in Computer Science
Cum laude in field with High Honors in Physics and Mathematics

EXPERIENCE

Cornell University, Cornell Tech New York, NY
Graduate Research Assistant Fall 2021 – Present
- Research in causal inference, machine learning, and robust data-driven decision-making. Adviser: Nathan Kallus.

Brookhaven National Laboratory Brookhaven, NY
Research Intern Summer 2024
- Developed causal inference methods for spatio-temporal applications in Earth Science. Mentor: Shinjae Yoo.

Netflix Los Gatos, CA
Machine Learning Intern Summer 2022
- Developed and built a causal machine learning model for quantifying the causal effect of watching a new title on long term user engagement. Mentors: Sudeep Das, Aish Fenton.

Microsoft Research Cambridge, MA
Senior Data and Applied Scientist 2019 – 2021
Data and Applied Scientist II 2017 – 2019
- Conducted research on machine learning-based causal inference techniques, contributing to top conference publications.
- Developed and published causal inference algorithms as a core contributor to the EconML library, supporting high-impact projects across various industries.
- Researched and improved subseasonal weather forecasting models, with results published in leading journals and conferences.

Microsoft Cambridge, MA
Software Development Engineer 2015 – 2017
- Developed and published scalable machine learning algorithms as a core contributor to MMLSpark, the Microsoft Machine Learning Library for Apache Spark.

Johns Hopkins University Baltimore, MD
Research Intern Summer 2014
- Developed clustering algorithms for clinical time series data to predict septic shock and created a dynamic web application for visualizing clusters and analyzing health data. Mentor: Suchi Saria.

SELECTED PUBLICATIONS

† - equal contribution, ‡ - alphabetical authors

[1] **Miruna Oprescu**, Jacob Dorn, Marah Ghoummaid, Andrew Jesson, Nathan Kallus, and Uri Shalit. B-learner: Quasi-oracle bounds on heterogeneous causal effects under hidden confounding. In *Proceedings of the 40th International Conference on Machine Learning*, pages 26599–26618. PMLR, 2023.

[2] Nathan Kallus† and **Miruna Oprescu**†. Robust and agnostic learning of conditional distributional treatment effects. In *International Conference on Artificial Intelligence and Statistics*, pages 6037–6060. PMLR, 2023.

[3] Keith Battocchi‡, Eleanor Dillon‡, Maggie Hei‡, Greg Lewis‡, **Miruna Oprescu**‡, and Vasilis Syrgkanis‡. Estimating the long-term effects of novel treatments. *Advances in Neural Information Processing Systems*, 34:2925–2935, 2021.

[4] **Miruna Oprescu**†, Vasilis Syrgkanis†, and Zhiwei Steven Wu†. Orthogonal random forest for causal inference. In *International Conference on Machine Learning*, pages 4932–4941. PMLR, 2019.

[5] Vasilis Syrgkanis, Victor Lei, **Miruna Oprescu**, Maggie Hei, Keith Battocchi, and Greg Lewis. Machine learning estimation of heterogeneous treatment effects with instruments. In *Advances in Neural Information Processing Systems*, pages 15193–15202, 2019. **Spotlight presentation.**

[6] **Miruna Oprescu**†, Vasilis Syrgkanis†, Keith Battocchi†, Maggie Hei†, and Greg Lewis†. EconML: A Machine Learning Library for Estimating Heterogeneous Treatment Effects. In *CausalML Workshop, NeurIPS*, 2019. **Spotlight presentation.**

TALKS

Uncertainty Quantification in Causal Inference: Sharp and Efficient Bounds on Heterogeneous Causal Effects Under Hidden Confounding
Computational Science Seminar, Brookhaven National Laboratory, 2023. Invited talk.

Causal Inference and Machine Learning in Practice with EconML and CausalML: Industrial Use Cases at Microsoft, TripAdvisor, Uber
The SIGKDD Conference on Knowledge Discovery & Data Mining, 2021. Accepted talk.

EconML: A Machine Learning Library for Estimating Heterogeneous Treatment Effects
Open Data Science Conference East, 2019. Invited Talk.

MMLSpark: Lessons from Building a SparkML Compatible Machine Learning Library
Spark Summit Europe, 2017. Accepted talk.

HONORS & AWARDS

Department of Energy Computational Science Graduate Fellowship	2022 – 2026
Meta PhD Research Fellowship Finalist	2022
cum laude, Harvard University	2015
High Honors, Harvard University Physics Department	2015
Derek C. Bok Award for Distinction in Teaching (*Data Science*), Harvard	2014

SERVICE

Peer Reviewer

- Conference on Neural Information Processing Systems (NeurIPS) 2021-2024
- International Conference on Machine Learning (ICML) 2024
- International Conference on Artificial Intelligence and Statistics (AISTATS) 2024

TEACHING

Teaching Assistant — Cornell University

- Learning, Inference, and Decision Making from Data — Spring 2022
- Applied Machine Learning — Fall 2021

Teaching Fellow — Harvard University

- Mechanics and Special Relativity — Fall 2014
- Data Science — Fall 2014
- Linear Algebra and Real Analysis — Spring 2013
- Algebra I — Fall 2013

PUBLICATIONS FULL LIST

† - equal contribution, ‡ - alphabetical authors

Latest publications available on Google Scholar.

PREPRINTS

[1] **Miruna Oprescu** and Nathan Kallus. Estimating heterogeneous treatment effects by combining weak instruments and observational data. *arXiv preprint arXiv:2406.06452*, 2024.

[2] Andrew Bennett‡, Nathan Kallus‡, **Miruna Oprescu**‡, Wen Sun‡, and Kaiwen Wang‡. Efficient and sharp off-policy evaluation in robust markov decision processes. *arXiv preprint arXiv:2404.00099*, 2024.

CONFERENCE PUBLICATIONS

[1] Andrew Bennett‡, Nathan Kallus‡, and **Miruna Oprescu**‡. Low-rank mdps with continuous action spaces. In *International Conference on Artificial Intelligence and Statistics*, pages 4069–4077. PMLR, 2024.

[2] **Miruna Oprescu**, Jacob Dorn, Marah Ghoummaid, Andrew Jesson, Nathan Kallus, and Uri Shalit. B-learner: Quasi-oracle bounds on heterogeneous causal effects under hidden confounding. In *Proceedings of the 40th International Conference on Machine Learning*, pages 26599–26618. PMLR, 2023.

[3] Nathan Kallus† and **Miruna Oprescu**†. Robust and agnostic learning of conditional distributional treatment effects. In *International Conference on Artificial Intelligence and Statistics*, pages 6037–6060. PMLR, 2023.

[4] Soukayna Mouatadid, Paulo Orenstein, Genevieve Flaspohler, Judah Cohen, **Miruna Oprescu**, Ernest Fraenkel, and Lester Mackey. Adaptive bias correction for improved subseasonal forecasting. *Nature Communications*, 14(1):3482, 2023.

[5] Keith Battocchi‡, Eleanor Dillon‡, Maggie Hei‡, Greg Lewis‡, **Miruna Oprescu**‡, and Vasilis Syrgkanis‡. Estimating the long-term effects of novel treatments. *Advances in Neural Information Processing Systems*, 34:2925–2935, 2021.

[6] Genevieve E Flaspohler, Francesco Orabona, Judah Cohen, Soukayna Mouatadid, **Miruna Oprescu**, Paulo Orenstein, and Lester Mackey. Online learning with optimism and delay. In *International Conference on Machine Learning*, pages 3363–3373. PMLR, 2021.

[7] **Miruna Oprescu**†, Vasilis Syrgkanis†, and Zhiwei Steven Wu†. Orthogonal random forest for causal inference. In *International Conference on Machine Learning*, pages 4932–4941. PMLR, 2019.

[8] Vasilis Syrgkanis, Victor Lei, **Miruna Oprescu**, Maggie Hei, Keith Battocchi, and Greg Lewis. Machine learning estimation of heterogeneous treatment effects with instruments. In *Advances in Neural Information Processing Systems*, pages 15193–15202, 2019. **Spotlight presentation.**

[9] **Miruna Oprescu**[†], Vasilis Syrgkanis[†], Keith Battocchi[†], Maggie Hei[†], and Greg Lewis[†]. EconML: A Machine Learning Library for Estimating Heterogeneous Treatment Effects. In *CausalML Workshop, NeurIPS*, 2019. **Spotlight presentation.**

[10] K Arbour, **M Oprescu**, J Hakim, H Rizvi, M Leiserson, M Ginsburg, A Plodkowski, J Sauter, I Preeshagul, S Gillett, et al. Multifactorial Model to Predict Response to PD-(L) 1 Blockade in Patients with High PD-L1 Metastatic Non-Small Cell Lung Cancer. *Journal of Thoracic Oncology*, 14(10):S290, 2019.