

David Madras

<http://www.cs.toronto.edu/~madras/>
madras@cs.toronto.edu

RESEARCH INTERESTS

My research focuses on building more trustworthy machine learning systems, primarily through improving robustness to spurious correlations and mitigating unfairness and discrimination. I aim to do research which acknowledges the role of machine learning models in larger systems, and am interested in creating more systematic evaluation procedures, understanding measurement error in data collection, and the role of automated tools in decision-making. I am also interested in causal modelling, calibration, uncertainty estimation, and representation learning.

EDUCATION

- JAN. 2018 - PRESENT PH.D. Candidate (expected May 2022), **University of Toronto**, Toronto, Canada
Advisor: [Richard Zemel](#)
- SEP. 2016 - JAN. 2018 MASTER OF SCIENCE, **University of Toronto**, Toronto, Canada
Thesis: Predict Responsibly: Improving Fairness and Accuracy by Learning to Defer
Advisor: [Richard Zemel](#)
- SEP. 2011 - JUN. 2016 BACHELOR OF SCIENCE, **University of Toronto**, Toronto, Canada
Major: Computer Science (specialist), with minors in Mathematics and Statistics
GPA: **3.92** (Dean's List)

PUBLICATIONS & PREPRINTS

- OCT. 2021 [Madras D.](#), Zemel R. [Identifying and Benchmarking Natural Out-of-Context Prediction Problems](#). *Neural Information Processing Systems, 2021*.
- SEP. 2020 Loewe S.*, [Madras D.*](#), Zemel R., Welling M. [Amortized Causal Discovery: Learning to Infer Causal Graphs from Time-Series Data](#). *Causal Learning and Reasoning (CLearR), 2022*.
- JUL. 2020 Creager E., [Madras D.](#), Pitassi T., Zemel R. [Causal Modelling for Fairness in Dynamical Systems](#). *International Conference on Machine Learning, 2020*.
- APR. 2020 [Madras D.](#), Atwood J., D'Amour, A. [Detecting Underspecification with Local Ensembles](#). *International Conference on Learning Representations, 2020*.
- APR. 2020 McCoy L., Burkell J., Card D., Davis B., Gichoya J., Le Page S., [Madras D.](#) [Beyond "In the Loop": On The Role of Meaningful Human Control in High-Stakes Machine-Human Partnerships](#). (*Oral*), *WeRobot 2020*.
- JUN. 2019 Creager E., [Madras D.](#), Jacobsen J.-H., Weis M., Pitassi T., Zemel R. [Flexibly Fair Representation Learning by Disentanglement](#). *International Conference on Machine Learning, 2019*.
- JAN. 2019 [Madras D.](#), Creager E., Pitassi T., Zemel R. [Fairness Through Causal Awareness: Learning Latent-Variable Models for Biased Data](#). *ACM Conference on Fairness, Accountability and Transparency (ACM FAT*), 2019*.
- DEC. 2018 [Madras D.](#), Pitassi T., Zemel R. [Predict Responsibly: Improving Fairness and Accuracy by Learning to Defer](#). *Neural Information Processing Systems, 2018*.
- JUN. 2018 [Madras D.*](#), Creager E.*, Pitassi T., Zemel R. [Learning Adversarially Fair and Transferable Representations](#). *International Conference on Machine Learning, 2018*.
- FEB. 2018 Chan T.C.Y., [Madras D.*](#), Puterman, M. [Improving fairness in match play golf through enhanced handicap allocation](#). *Journal of Sports Analytics, 2018*.
- FEB. 2017 Allen, S.*, [Madras D.*](#), Ye Y.*, Zanotti, G.* [Change-point Detection Methods for Body-Worn Video](#). *Society for Industrial and Applied Mathematics (SIAM) Undergraduate Research Online, 2018*.

WORKSHOP PAPERS

- DEC. 2021 **Madras D.**, Zemel R. [Understanding Post-hoc Adaptation for Improving Subgroup Robustness](#). *Distribution Shifts: Connecting Methods and Applications, NeurIPS 2021*.
- DEC. 2021 **Madras D.**, Zemel R. Towards Systematic Evaluation in Machine Learning through Automated Stress Test Creation. *Data-Centric AI Workshop, NeurIPS 2021*.
- DEC. 2020 Adragna R., Creager E., **Madras D.**, Zemel R. [Fairness and Robustness in Invariant Learning: A Case Study in Toxicity Classification](#). *Algorithmic Fairness through the Lens of Causality and Interpretability (Oral), NeurIPS 2020*.
- APR. 2020 Creager E., **Madras D.**, Pitassi T., Zemel R. Causal Modeling for Fairness in Dynamical Systems: A Case Study in Lending. *Workshop on Causal Learning and Decision-Making, ICLR 2020*.
- JUN. 2019 **Madras D.**, Atwood J., D'Amour, A. Detecting Extrapolation with Influence Functions. *Workshop on Uncertainty and Robustness in Deep Learning (Oral), ICML 2019*.
- DEC. 2018 **Madras D.**, Creager E., Pitassi T., Zemel R. Fairness Through Causal Awareness: Learning Latent-Variable Models for Biased Data. *Workshop on Ethical, Social and Governance Issues in AI (Spotlight), NeurIPS 2018*.
- DEC. 2017 **Madras D.**, Pitassi T., Zemel R. Predict Responsibly: Improving Fairness and Accuracy by Learning to Defer. *Workshop on Transparent and Interpretable Machine Learning in Safety Critical Environments (Oral, Best Paper Award), NeurIPS 2017*.

AWARDS AND HONORS

- NOV. 2020 Schwartz Reisman Institute for Technology and Society Fellowship
- SEP. 2019 NSERC Alexander Graham Bell Canada Graduate Scholarship-Doctoral (CGS D)
- DEC. 2017 Best paper, NeurIPS 2017 Workshop on Transparent and Interpretable Machine Learning in Safety Critical Environments
- SEP. 2016 NSERC CGS Master's Fellowship
- MAR. 2016 2nd Place, Waterfront International Quantathon, Graduate Division
- MAY 2016 Research in Industrial Projects for Students scholarship, Institute for Applied Mathematics, UCLA
- MAY 2015 Kupcinet-Getz scholarship, Weizmann Institute of Science
- MAY 2014 NSERC Undergraduate Student Research Award
- SEP. 2014 Louis Savlov Scholarship, University of Toronto
- APR. 2013 Award of Excellence for Achievement in Computer Science – Department of Computer Science, University of Toronto
- SEP. 2011 Reuben Leonard Wells Scholarship

LEADERSHIP

- JUL. 2020 Co-organized inaugural workshop on [Participatory Approaches to Machine Learning](#) (ICML 2020)
- NOV. 2019 Organized inaugural [Pan-Canadian Self-Organizing Conference on Machine Learning](#) (co-chair)
- 2019-2020 Organized Causal Inference reading group (Vector Institute)
- 2018 Organized Fairness in Machine Learning reading group (Vector Institute)
- 2014-2015 Co-president of Undergraduate Artificial Intelligence Group (UAIG), University of Toronto

WORK EXPERIENCE

- FEB. 2019 **Research Intern**, Google Brain, Cambridge, MA, USA
- MAY. 2019 Worked on methods for unreliability detection and sensitivity analysis with the Google Brain team, hosted by Alex D'Amour and James Atwood.
- SEP. 2014 **Software Developer**, Ontario Financing Authority
- APR. 2015 Developed software using VB, C#, SQL; gathered requirements from traders and developers; succeeded in self-motivated work environment.

TEACHING EXPERIENCE

MAY 2019	African Institute of Mathematical Sciences , Kigali, Rwanda <i>Assistant Instructor</i> Three-week course on Fairness & Privacy in Machine Learning at the African Masters of Machine Intelligence (https://aimsammi.org/)
SEP. 2013 PRESENT	University of Toronto , Toronto, Canada <i>Teaching Assistant</i> CSC411: Machine Learning and Data Mining, CSC412: Probabilistic Learning and Reasoning, CSC108: Introduction to Computer Programming, CSC148: Introduction to Computer Science

SERVICE

REVIEWER: ICLR (2022, 2021, 2019), NeurIPS (2021, 2020, 2019), ICML (2019), FAccT (2021), AISTATS (2020), JMLR (2x)

MISCELLANY

OCT. 2019 Attended NBER Economics of Artificial Intelligence Conference & Young Scholars Workshop (invitation only)
MAY. 2019 Summer Institute on AI and Society, Alberta Machine Intelligence Institute
AUG. 2017 Attended Deep Learning Summer School, University of Montreal

GRADUATE COURSEWORK

WINTER 2020 Statistical Learning Theory
WINTER 2020 Algorithms for Collective Decision-Making
FALL 2018 Fundamentals of Cryptography
WINTER 2018 Computational Social Science
WINTER 2017 Visual Recognition with Text
WINTER 2017 Algorithms & Complexity in Private Data Analysis
FALL 2016 Inference & Generative Models
FALL 2016 Computational Linguistics

INVITED TALKS

AUG 2021 *Why Learn Fair Representations?*
KDD 2021 Deep Learning Day (Singapore)
SEP. 2020 *Fairness Through Causal Awareness*
HAI Reading Group, Apple (Seattle)
MAY 2020 *Causality in Machine Learning*
African Institute of Mathematical Sciences (Rwanda, Ghana)
AUG. 2019 *Machine Learning in Decision-Making Systems*
Schwartz Reisman Institute for Technology and Society (Toronto)
JUNE 2019 *Detecting Extrapolation with Influence Functions*
Workshop on Uncertainty and Robustness in Deep Learning, ICML 2019 (Long Beach)
MAY 2019 *Detecting Extrapolation with Influence Functions*
Google Brain (Cambridge)
FEB. 2019 *Predict Responsibly: Fairness in Machine Learning*
Princeton University (Princeton)
JAN. 2019 *Fairness Through Causal Awareness: Learning Latent-Variable Models for Biased Data*
ACM FAT* 2019 (Atlanta)
DEC. 2018 *Fairness Through Causal Awareness: Learning Latent-Variable Models for Biased Data*
Workshop on AI, Ethics & Governance, NeurIPS 2018 (Montreal)
OCT. 2018 *Predict Responsibly: Fairness in Machine Learning*
Toronto Deep Learning Series, Sidewalk Labs (Toronto)
SEP. 2018 *Learning Adversarially Fair and Transferable Representations*
Fairness in Machine Learning Workshop, Google Brain (Cambridge)
SEP. 2018 *Learning Adversarially Fair and Transferable Representations*
Google Brain (Toronto)
AUG. 2018 *Learning Adversarially Fair and Transferable Representations*
Radiation Medication Program Summer Series, Princess Margaret Hospital (Toronto)
JUL. 2018 *Learning Adversarially Fair and Transferable Representations*
Sunnybrook Health Sciences Centre (Toronto)

JUL. 2018 *Learning Adversarially Fair and Transferable Representations*
ICML 2018 (Stockholm)

APR. 2018 *Fairness in Machine Learning*
University in the Community (Toronto)

APR. 2018 *Predict Responsibly: Improving Fairness and Accuracy by Learning to Defer*
Design+AI Meetup, Normative AI (Toronto)

APR. 2018 *Predict Responsibly: Improving Fairness and Accuracy by Learning to Defer*
Workshop on Transparent and Interpretable ML
in Safety Critical Environments, NeurIPS 2017 (Long Beach)

MAR. 2018 *Panel Discussion: Fairness and Interpretability in Machine Learning*
Integrate AI (Toronto)

NOV. 2017 *Fairness in Machine Learning*
AI Day, University of Toronto (Toronto)