David Madras

madras@cs.toronto.edu
www.cs.toronto.edu/~madras

EDUCATION

PhD, Machine Learning Group, University of Toronto. *Expected: 2021.*
MSc, Machine Learning Group, University of Toronto. *Completed January 2018.*
BSc, University of Toronto. *Completed June 2016*
Major: Computer Science (specialist). Minors: Mathematics and Statistics. GPA: **3.92**

RESEARCH INTERESTS

Algorithmic fairness, causal inference, robustness to distributional shift, algorithmic decision-making, security/safety/privacy of AI systems, AI law/policy, generative modeling

PUBLICATIONS

Creager E., Madras D., Jacobsen J.-H., Weis M., Pitassi T., Zemel R. **Flexibly Fair Representation Learning by Disentanglement** *ICML, 2019.*
Madras D., Creager E., Pitassi T., Zemel R. **Fairness through Causal Awareness: Learning Causal Latent-Variable Models for Biased Data**. *ACM FAT*, 2019.*
Madras D., Creager E., Pitassi T., Zemel R. **Learning Adversarially Fair and Transferable Representations**. *ICML, 2018* ***(Oral, Long – 8.6% of submissions)***.
Madras D., Pitassi T., Zemel R. **Predict Responsibly: Increasing Fairness by Learning to Defer**. *NeurIPS 2018 & Neurips TIML Workshop 2017* ***(Oral, Best Paper Award).***
Chan T.C.Y., Madras D., Puterman M. **Improving Fairness in Match Play Golf through Enhanced Handicap Allocation**. *JSA, 2018.*
Allen S., Madras D., Ye Y., Zanotti G. **Change-point Detection Methods for Body-Worn Video**. *SIAM-URO, 2017.*

RESEARCH EXPERIENCE

Machine Learning Group, University of Toronto (current) *2016-present.*
Google Brain, Cambridge, USA (internship) *2019.*
Institute for Applied Mathematics, UCLA *2016.*
Dept. of Mechanical and Industrial Engineering, University of Toronto *2015-2016.*
Dept. of Computer Science & Applied Mathematics, Weizmann Institute of Science *2015.*
Computational Biology Group, University of Toronto *2014.*

GRADUATE COURSES

2018: Fundamentals of Cryptography, Computational Social Science
2017: Visual Recognition with Text, Algorithms & Complexity in Private Data Analysis
2016: Inference & Generative Models, Computational Linguistics

AWARDS

Best Paper, NIPS Transparent and Interpretable ML in Safety Critical Environments *2017.*
NSERC CGS Master's Grant *2016.*
2nd Place – Waterfront International Quantathon, Graduate Division *2016*

SERVICE

Reviewer: NeurIPS 2019, ICLR 2019 DebugML Workshop, ICML 2019

Economics of AI Workshop: Application Essay

David Madras

I am a third-year PhD student in the Machine Learning (ML) group at the University of Toronto. Over the past three years, I have done research on a range of ML topics. My research focus has been on the nascent field of *fairness* in machine learning. I have been primarily interested in understanding decision-making systems, both automated and human. For instance, in my paper "Predict Responsibly: Improving Fairness and Accuracy by Learning to Defer" (NeurIPS 2018), we explore the algorithmic ramifications of human decision-makers using machine learning systems as assistive tools. My work also explores concepts of fairness in ML by applying state-of-the-art work in deep generative models (an important field in modern ML research) for tasks such as censoring sensitive information or decision-making from historically biased datasets, with the goal of building maximally flexible and scalable models.

Recently, I have become interested in economics and AI through an interest in causal inference (CI). CI tools seem to answer many essential decision-making questions which are left unanswered by ML. In my FAT* 2019 paper "Fairness through Causal Awareness", we explore a method for CI under hidden confounding, an area of paramount practical importance. Additionally, I just finished an internship at Google working with a CI researcher, where we developed a method inspired by robust statistics with the goal of verifying applicability of CI methods, and its applicability to deep learning models (ICML workshop paper forthcoming). I believe that the interaction of ML and CI is of great importance for harnessing large datasets and flexible models to make good decisions. The Economics and AI workshop seems like the ideal place to learn more about this intersection, and to meet other researchers interested in these topics.