

LEONARDO COTTA

✉ leonardo.cotta@vectorinstitute.ai

🌐 <https://cottascience.github.io/>

📍 Toronto, ON, Canada

MACHINE LEARNING RESEARCHER

I develop machine learning and causal inference methods for complex discrete data structures, including sequences and graphs. I am particularly inspired by applications in biochemistry and medicine.

EDUCATION

Doctor of Philosophy in Computer Science Purdue University, West Lafayette, IN, USA Advisor: Bruno Ribeiro	2017 – 2022
Bachelor of Science in Computer Science Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil	2012 – 2016

RESEARCH POSITIONS

Vector Institute, Toronto, ON, Canada Distinguished Postdoctoral Fellow hosted by Chris J. Maddison	2022 – present
Purdue University, West Lafayette, IN, USA Graduate Research Assistant	2020 – 2022
Intel Labs, Santa Clara, CA, USA AI Research Intern	May-August 2021
Purdue University, West Lafayette, IN, USA Graduate Teaching Assistant	2019 – 2020
Purdue University, West Lafayette, IN, USA Graduate Research Assistant	2018 – 2019
Purdue University, West Lafayette, IN, USA Qatar Computing Research Institute Fellow	2017 – 2018
Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil Undergraduate Research Assistant	2016 – 2017
LG Electronics, Belo Horizonte, MG, Brazil Undergraduate Research Assistant	2015

AWARDS AND HONORS

Distinguished Postdoctoral Fellowship, Vector Institute (Canada)	2022
Graduate Research Fellowship, Qatar Computing Research Institute (USA)	2017
Science Without Borders Scholarship, University of Calgary/CNPq (Brazil)	2014
Young Talents of Science Award, CAPES (Brazil)	2012

SERVICE

Since 2021, I am a regular reviewer for NeurIPS, ICML, and ICLR.

Area chair for AAAI 2022.

Ad-hoc reviewer for Journal of Computational and Graphical Statistics.

Organizer of LoG (Learning on Graphs) 2022.

Mentor at the LOGGML Summer School in 2022.

TEACHING

Cross-Training in AI and Laboratory Knowledge for Drug Discovery, University of Toronto 2025
Tutor — ≈ 30 students.

Computational Methods in Optimization, CS520, Purdue University 2020
Teaching Assistant — ≈ 30 students.

Introduction to Data Science, CS242, Purdue University 2019
Head Teaching assistant — ≈ 200 students.

REFERENCES

- Prof. Chris J. Maddison, University of Toronto cmaddis@cs.toronto.edu
- Prof. Bruno Ribeiro, Purdue University ribeirob@purdue.edu
- Prof. Rahul Krishnan, University of Toronto rahulgk@cs.toronto.edu
- Prof. Christopher Morris, RWTH Aachen University morris@cs.rwth-aachen.de

SELECTED PUBLICATIONS

For full list and metrics, see [Google Scholar](#).

PREPRINTS

- Test-Time Fairness and Robustness in Large Language Models

L Cotta, CJ Maddison

arXiv preprint, 2024

- Boosting the Predictive Power of Protein Representations with a Corpus of Text Annotations

H Duan, M Skreta, L Cotta, EM Rajaonson, N Dhawan, A Aspuru-Guzik, CJ Maddison

arXiv preprint, 2024

PUBLISHED

- End-To-End Causal Effect Estimation from Unstructured Natural Language Data

N Dhawan, L Cotta, K Ullrich, RG Krishnan, CJ Maddison

NeurIPS, 2024

- Probabilistic Invariant Learning with Randomized Linear Classifiers

L Cotta, G Yehuda, A Schuster, CJ Maddison

NeurIPS, 2023

- Causal Lifting and Link Prediction

L Cotta, B Bevilacqua, N Ahmed, B Ribeiro

Proceedings of the Royal Society A, 2023

- Reconstruction for Powerful Graph Representations

L Cotta, C Morris, B Ribeiro

NeurIPS, 2021

- Unsupervised Joint k -node Graph Representations with Compositional Energy-Based Models

L Cotta, CHC Teixeira, A Swami, B Ribeiro, 2020

NeurIPS, 2020

- Graph Pattern Mining and Learning through User-defined Relations

CHC Teixeira, L Cotta, B Ribeiro, W Meira Jr

IEEE ICDM, 2018

- Understanding the role of mobility in real mobile ad-hoc networks connectivity

L Cotta, POS Vaz de Melo, AAF Loureiro

IEEE ISCC, 2017

- AoT: Authentication and Access Control for the Entire IoT Device Life-Cycle

AL Maia Neto, ALF Souza, Í Cunha, M Nogueira, IO Nunes, L Cotta, N Gentile, AAF Loureiro, DF Aranha, HK Patil, LB Oliveira

ACM Sensys, 2016

- Using FIFA Soccer video game data for soccer analytics

L Cotta, POS Vaz de Melo, F Benevenuto, AAF Loureiro

Workshop on Large Scale Sports Analytics (KDD), 2016

- Nomadikey: User authentication for smart devices based on nomadic keys

L Cotta, AL Fernandes, LTC Melo, LFZ Saggioro, F Martins, ALM Neto, AAF Loureiro, Í Cunha, LB Oliveira

IEEE ICC, 2016

INVITED TALKS

- Structure and scale: A tale of two dogmas in AI4Science, 2025

Broad Institute, Arc Institute

- Test-time algorithms for language models, 2024

Deep Learning & Reinforcement Learning Summer School, Toronto

- Causal Lifting and Link Prediction, 2023

Johns Hopkins, CISS

- Higher-order reasoning with graph data, 2022

Vector Institute, Microsoft Research

LANGUAGES

English, Fluent

Portuguese, Fluent (Native)

Spanish, Basic

CITIZENSHIP

Brazilian

Italian