LEONARDO COTTA

MACHINE LEARNING RESEARCHER

leonardo.cotta@vectorinstitute.ai
https://cottascience.github.io/
London/Oxford, UK

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 — 2017 – 2022
Purdue University, West Lafayette, IN, USA
Advisor: Bruno Ribeiro

Bachelor of Science in Computer Science — 2012 – 2016
Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

RESEARCH POSITIONS

Ellison Institute of Technology, Oxford, UK — 2025 – Present
Research Scientist

Vector Institute, Toronto, ON, Canada — 2022 – 2025
Distinguished Postdoctoral Fellow hosted by Chris J. Maddison

Purdue University, West Lafayette, IN, USA — 2020 – 2022
Graduate Research Assistant

Intel Labs, Santa Clara, CA, USA — May-August 2021
AI Research Intern

Purdue University, West Lafayette, IN, USA — 2019 – 2020
Graduate Teaching Assistant

Purdue University, West Lafayette, IN, USA — 2018 – 2019
Graduate Research Assistant

Purdue University, West Lafayette, IN, USA — 2017 – 2018
Qatar Computing Research Institute Fellow

Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil — 2016 – 2017
Undergraduate Research Assistant

LG Electronics, Belo Horizonte, MG, Brazil — 2015
Undergraduate Research Assistant

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

- Danilo J Rezende, Ellison Institute of Technology drezende@eit.org
- Chris J Maddison, University of Toronto cmaddis@cs.toronto.edu
- Bruno Ribeiro, Purdue University ribeirob@purdue.edu

SELECTED PUBLICATIONS

For full list and metrics, see Google Scholar.

- Bayesian Sensitivity of Causal Inference Estimators under Evidence-Based Priors
N Dhawan, D Shen, L Cotta, CJ Maddison
Transactions in Machine Learning Research, 2025

- Measuring Scientific Capabilities of Language Models with a Systems Biology Dry Lab
H Duan, SZ Lu, CF Harrigan, N Desai, J Lu, M Koziarski, L Cotta, CJ Maddison
NeurIPS, 2025

- Test-Time Fairness and Robustness in Large Language Models
L Cotta, CJ Maddison
Transactions in Machine Learning Research, 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
Nature Machine Intelligence, 2025

- 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 Gentille, 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)

CITIZENSHIP

Brazilian

Italian