

Lucio Dery

GHC 5101, 5000 Forbes Ave, Pittsburgh PA 15213

🌐 [ldery.github.io](https://github.com/ldery)

✉ ldery@andrew.cmu.edu

🌐 <https://github.com/ldery>

EDUCATION

Carnegie Mellon University

PhD in Computer Science

Pittsburgh, PA

Expected Graduation: 2024

Stanford University

*MS in Computer Science ** Tau Beta Pi*

Stanford, CA

Sept 2016 - June 2018

Stanford University

*BS in Physics + Minor in Computer Science ** With Distinction*

Stanford, CA

Sept 2013 - June 2018

RESEARCH INTERESTS

- Transfer Learning, Meta-Learning, Multitasking, Weak Supervision, Natural Language Processing

PUBLICATIONS / TALKS

Peer Reviewed Publications.....

- *Cross-Modal Fine-Tuning: Align then Refine*

Junhong Shen, Liam Li, **Lucio M. Dery**, Corey Staten, Mikhail Khodak, Graham Neubig, Ameet Talwalkar

ICML 2023 ([Oral](#)) [[Paper](#)][[Code](#)]

- *AANG: Automating Auxiliary Learning*

Lucio M. Dery, Paul Michel, Mikhail Khodak, Graham Neubig, Ameet Talwalkar.

ICLR, 2023 ([Spotlight](#)) [[Paper](#)] [[Code](#)]

- *Multistep planning for automated hyper-parameter optimization : An exploration via OptFormer*

Lucio M. Dery, Yutian Chen, Abram L Friesen, Marc'Aurelio Ranzato.

[FMDM Workshop - NeurIPS 2022](#) [[Paper](#)]

- *Should We Be Pre-training? An Argument for End-task Aware Training as an Alternative*

Lucio M. Dery, Paul Michel, Ameet Talwalkar, Graham Neubig.

ICLR, 2022 [[Paper](#)][[Code](#)]

- *Auxiliary task update decomposition: the good, the bad and the neutral*

Lucio M. Dery, Yann Dauphin, David Grangier.

ICLR, 2021 [[Paper](#)][[Code](#)]

- *Audio to Body Dynamics*

Eli Shlizerman, **Lucio M. Dery**, Hayden Schoen, Ira Kemelmacher.

CVPR, 2018 [[Paper](#)][[Code](#)][[Press](#)]

- *Finding 'It': Weakly-Supervised Reference-Aware Visual Grounding in Instructional Video*

D.A-Huang, Shyamal Buch, **Lucio M. Dery**, Animesh Garg, Li Fei-Fei, Juan Carlos Niebles.

CVPR, 2018 ([Oral](#)) [[Paper](#)][[Code](#)]

- *Weakly supervised classification in high energy physics*

Lucio M. Dery, Benjamin Nachman, Francesco Rubbo, Ariel Schwartzman

Journal of High Energy Physics 2017.5 (2017): 1-11 [[Paper](#)] [[Code](#)]

Patents.....

- o **Lucio M. Dery**, Yann Dauphin, David Grangier. “Training Neural Networks Using Auxiliary Task Update Decomposition” [[Patent](#)]

Invited Talks.....

- o *AANG: Automating Auxiliary Learning* [ICLR - 2023]
- o *Auxiliary Task Update Decomposition* [ACMI Lab - 2021]
- o *Audio to Body Dynamics* [Black In A.I Workshop @ NeurIPS - 2018]
- o *Weakly Supervised Classification in High Energy Physics* [[ACAT](#) - 2017]

INDUSTRY EXPERIENCE

Apple - AI/ML

Cupertino, CA

AI/ML Research Intern

May 2023 - August 2023

- o Structured pruning of pre-trained language models under limited target data.

Google Research

Remote

Student Researcher

March 2023 - May 2023

DeepMind

London, UK

Research Scientist Intern

May 2022 - August 2022

- o Worked on data-driven hyper-parameter optimization using transformers ([OptFormer](#))
- o Focused on inference-time improvements to OptFormer
- o **Artifacts:** [[Paper](#)]

Google Brain - Google

Remote

Research Scientist Intern

June 2020 - August 2020

- o Leveraged out-of-distribution data via Gradient Alignment
- o Examined auxiliary task gradients within subspace spanned by primary task gradients
- o **Artifacts:** [[Paper](#)][[Code](#)]

Facebook A.I Research - Facebook

Seattle, WA

Research Engineer

July 2018 – July 2019

- o Studied learning Neural Knowledge Graphs by Generating Wikipedia
- o Probed Commonsense and World Knowledge Capabilities of State-of-the-Art Co-reference Models
- o **Artifacts:** [[Open-sourced Audio to Body Dynamics](#)], [[Contributed to FAIRSEQ](#)]

Applied Machine Learning - Facebook

Seattle, WA

Software Engineering Intern

June 2017 – August 2017

- o Worked on Audio-Visio Multimodal Learning for understanding human mannerisms
- o Developed recurrent architecture for learning transformations from audio features to body key-points
- o **Artifacts:** [[Paper](#)][[Code](#)][[Press](#)]

Terra Bella - Google

Mountain View, CA

Software Engineering Intern

June 2016 – August 2016

- o Applied unsupervised learning techniques to Satellite images to cluster similar socio-economic regions and detect changing regions over time
- o Extensive feature engineering through experimentation with remote sensing signal spaces like NDVI (Normalized Difference of Vegetation Index), MSAVI and NDBI
- o Built Tensor Flow model that utilized Inception V3 featurization of remote sensing signal spaces to automatically identify similar regions like Golf Courses or Airports within and across cities

Google Analytics - Google

Mountain View, CA

Engineering Practicum Intern

June 2015 – August 2015

- Conducted background experimentation and comparative performance visualizations in R on time series prediction algorithms in Analytics libraries against third party algorithms
- Implemented Autoregressive Integrated Moving Averages (ARIMA) time series forecasting algorithm. Resulting implementation was on average faster than R implementation and of comparable accuracy
- Exposed ensemble mode API that allows developers to use suite of forecasting algorithms

TEACHING EXPERIENCE

- Teaching Assistant, [Advanced Natural Language Processing](#), Fall 2022
- Section Leader, [Stanford Code In Place](#), Spring 2020
- Computer Vision Instructor, [African Masters in Machine Intelligence](#), Summer 2019
- Head Teaching Assistant, Deep Learning (CS230) Stanford University, Spring 2018
- Course Assistant, Deep Learning (CS230), Stanford University, Winter 2018
- Course Assistant, Machine Learning (CS229), Stanford University, Autumn 2017
- Section Leader, Programming Methodology (CS106A), Stanford University, 2014 - 2017
- Section Leader, Programming Abstractions (CS106B), Stanford University, 2014 - 2017
- Summer School Instructor, [Enza Academy](#), Summer 2015

HONORS / AWARDS

- [2nd Place Two Sigma Diversity PhD Fellowship](#)
- Stanford Chapter Tau Beta Pi Honor Society
- University Distinction, top 15% of graduating class, Stanford University
- Stanford Black Community Center Award for Academic Excellence
- Stanford Center for African Studies Leadership and Service Award
- Stanford Computer Science Department TA Award
(\$1000 awarded to top 5% of Course Assistants in Spring 2018)
- 3rd Best Student, West African Senior Secondary Certificate Examination
(out of over 2.1 million students from Anglophone West Africa in 2013)[[Press](#)]
- 2nd Place, Ghana National Math and Science Olympiad (out of 32 Selected Schools) [[Press](#)]

SERVICE

- Reviewer
 - ICLR 2022,2023
 - ICML 2022,2023
 - NeurIPS 2022,2023
 - TMLR 2023
- Volunteer ICLR 2023
- Graduate School Application Mentorship - [Black In AI](#)
- Social Chair - CMU African Grad Student Association
- Mentorship Assistant Lead - CMU CSD PhD Student Council
- Mentor & Mock Interviewer - Underrepresented minorities seeking Software Engineering Roles