

Lucio Dery

GHC 5101, 5000 Forbes Ave, Pittsburgh PA 15213

🌐 ldery.github.io

✉ ldery@andrew.cmu.edu

🔗 <https://github.com/ldery>

EDUCATION

Carnegie Mellon University

PhD in Computer Science

Pittsburgh, PA

Expected Graduation: June 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

Papers.....

1. **Lucio M. Dery**, Yutian Chen, Abram L Friesen, Marc'Aurelio Ranzato. "Multistep planning for automated hyper-parameter optimization : An exploration via OptFormer" [\[Workshop on Foundation Models For Decision Making - NeurIPS 2022\]](#) [\[Paper\]](#)
2. **Lucio M. Dery**, Paul Michel, Ameeta Talwalkar, Graham Neubig. "Should We Be Pre-training? An Argument for End-task Aware Training as an Alternative" ICLR, 2022 [\[Paper\]](#)[\[Code\]](#)
3. **Lucio M. Dery**, Yann Dauphin, David Grangier. "Auxiliary task update decomposition: the good, the bad and the neutral". ICLR, 2021 [\[Paper\]](#)[\[Code\]](#)
4. Eli Shlizerman, **Lucio M. Dery**, Hayden Schoen, Ira Kemelmacher. "Audio to Body Dynamics." CVPR, 2018 [\[Paper\]](#)[\[Code\]](#)[\[Press\]](#)
5. D.A-Huang, Shyamal Buch, **Lucio M. Dery**, Animesh Garg, Li Fei-Fei, Juan Carlos Niebles. "Finding 'It': Weakly-Supervised Reference-Aware Visual Grounding in Instructional Video." CVPR, 2018 [\(Oral\)](#)[\[Paper\]](#)[\[Code\]](#)
6. **Lucio M. Dery**, et al. "Weakly supervised classification in high energy physics." Journal of High Energy Physics 2017.5 (2017): 1-11 [\[Paper\]](#) [\[Code\]](#)

Pre-prints / Ongoing Work.....

1. **Lucio M. Dery**, Paul Michel, Mikhail Khodak, Graham Neubig, Ameeta Talwalkar. "AANG: Automating Auxiliary Learning" (under submission) [\[Paper\]](#)
2. Junhong Shen, Liam Li, **Lucio M. Dery**, Corey Staten, Mikhail Khodak, Graham Neubig, Ameeta Talwalkar. "A Practical Framework for Cross-Domain Transfer Learning" (under submission)

Patents.....

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

Invited Talks.....

- *Auxiliary Task Update Decomposition*, [ACMI Lab - 2021]
- *Audio to Body Dynamics* [Black In A.I Workshop @ NeurIPS - 2018]
- *Weakly Supervised Classification in High Energy Physics* [International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT) - 2017]

INDUSTRY EXPERIENCE

DeepMind

London, UK

Research Scientist Intern

May 2022 - August 2022

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

Google Brain - Google

Remote

Research Scientist Intern

June 2020 - August 2020

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

Facebook A.I Research - Facebook

Seattle, WA

Research Engineer

July 2018 - July 2019

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

Applied Machine Learning - Facebook

Seattle, WA

Software Engineering Intern

June 2017 - August 2017

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

Terra Bella - Google

Mountain View, CA

Software Engineering Intern

June 2016 - August 2016

- Applied unsupervised learning techniques to Satellite images to cluster similar socio-economic regions and detect changing regions over time
- Extensive feature engineering through experimentation with remote sensing signal spaces like NDVI (Normalized Difference of Vegetation Index), MSAVI and NDBI
- 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
- 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, NeurIPS 2022
- Graduate School Application Mentorship - [Black In AI](#)
- Mock Interviewer - Underrepresented minorities seeking Software Engineering Roles