CALEB N. ELLINGTON

calebellington.com

Office 7405, Gates-Hillman Center
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

PUBLICATIONS & PRESENTATIONS

Ellington, Caleb, Ben Lengerich, Thomas B.K. Watkins, Jiekun Yang, Manolis Kellis, Eric P. Xing. "**Sample-Specific Contextualized Graphical Models Using Clinical and Molecular Data Reveal Transcriptional Network Heterogeneity Across 7000 Tumors.**"

- Platform Talk, Cold Spring Harbor Laboratory – Biological Data Science, 2022.
- Poster Presentation, Machine Learning for Health, 2022.
- Poster Presentation, Graph Learning for Industrial Applications NeurIPS Workshop, 2022.

Lengerich, Ben, Caleb Ellington, Bryon Aragam, Eric P. Xing, and Manolis Kellis. "**NOTMAD: Estimating Bayesian Networks with Sample-Specific Structures and Parameters.**" ArXiv:2111.01104 [Cs, Stat], November 1, 2021. http://arxiv.org/abs/2111.01104.

Chen, Xinshi, Haoran Sun, Caleb Ellington, Eric Xing, and Le Song. "**Multi-Task Learning of Order-Consistent Causal Graphs.**" In Advances in Neural Information Processing Systems, 34:11083–95. Curran Associates, Inc., 2021. https://proceedings.neurips.cc/paper/2021/hash/5c3a3b139a11689e0bc55abd95e20e39-Abstract.html.

Lengerich, Benjamin J, Mark E Nunnally, Yin Aphinyanaphongs, Caleb Ellington, and Rich Caruana. "**Automated Interpretable Discovery of Heterogeneous Treatment Effectiveness: A COVID-19 Case Study.**" J. Biomed. Inform., April 30, 2022, 104086. https://doi.org/10.1016/j.jbi.2022.104086.

SOFTWARE

I created contextualized.ml (repo: https://github.com/cnellington/Contextualized) as an sklearn-style machine learning toolbox to help informaticians, statisticians, and ML researchers across domains infer models, distributions, and functions with context-specific parameters.

POSITIONS

- **Ph.D. Student in Computational Biology** — Sept 2020 - Present
  Statistical Artificial Intelligence and Integrative Genomics (SAILING) Lab, Carnegie Mellon University — Pittsburgh, PA
- **Research Assistant** — Jan 2019 - June 2020
  Baker Lab, Institute for Protein Design, University of Washington — Seattle, WA
- **Software Engineer Intern** — June 2019 - Aug 2019
  Indeed.com — Austin, TX
- **Research Assistant** — Jan 2017 - Dec 2018
  Klavins Lab, Automated Bio-Fabrication Facility, University of Washington — Seattle, WA
- **Biomedical Engineering Consultant** — Aug 2018 - Sept 2018
  Dhulikhel Hospital / Bioengineering Department, University of Washington — Dhulikhel, Nepal
- **Software Engineer Intern** — June 2018 - Aug 2018
  Amazon.com — Seattle, WA
- **Software Engineer Intern** — June 2017 - Aug 2017
  Inscripta Biotechnology — Boulder, CO

EDUCATION

Carnegie Mellon University Sept 2020 - Present
Ph.D. Student in Computational Biology Pittsburgh, PA
Advised by Eric P. Xing
Coursework: Probabilistic Graphical Models, Machine Learning, Intermediate Statistics, Convex Optimization, Computational Genomics, Advanced Genetics, Computational Structural Biology, Cellular Systems Biology, Lab Methods

University of Washington Sept 2016 - Jun 2020
B.Sc. Computer Science Seattle, WA
B.Sc. Bioengineering (Honors)

AWARDS

- **GRFP Honorable Mention (NSF)** Apr 2020
  Project: *Deep Generative Models for Ligand-based Protein Design*
- **Fulbright Semifinalist (Fulbright)** Apr 2020
  Project: *Freelance Ambulances: Integrating Nepal's private transportation with emergency services*
- **Levinson Emerging Scholar (Washington Research Foundation)** Sept 2019
  Awarded to highly motivated students demonstrating exceptional communication skills pursuing advanced bioscience careers.
- **Husky 100 Awardee (University of Washington)** Mar, 2019
  The Husky 100 recognizes 100 students (graduate and undergraduate) from all UW campuses who are making the most of their time at UW.
- **Mary Gates Research Scholar (University of Washington)** Feb, 2018
  - Project: *Data-Driven Genetic Engineering with Hydra vulgaris*
  - Awarded for strong academic merit, good mentor relationships, intensity of research experience, and long-term impact of the proposed project.

TEACHING & MENTORING

Teaching Assistant, 10-701 Ph.D. Introduction to Machine Learning Jan 2022 - May 2022
Instructors: Henry Chai, Maria Balkan Carnegie Mellon University

Teaching Assistant, 02-319 Genetics and Epigenetics of the Brain Aug 2021 - Dec 2021
Instructors: Andreas Pfenning Carnegie Mellon University

Current Mentees
- Alyssa Lee, Computational Biology BS Student, CMU
- Jannik Deuschel, Machine Learning MS Student, KIT
- Ding Bai, Machine Learning Ph.D. Student, MBZUAI
- Tianjun Yao, Machine Learning Ph.D. Student, MBZUAI
- Juwayni Lucman, Artificial Intelligence MS Student, MBZUAI
- Wesley Lo, Mathematics BS Student, MIT

COMMUNITY INVOLVEMENT

- **Admissions Committee**, CMU-Pitt Computational Biology Ph.D. Program January, 2023
- **Junior Chair**, Round-table: Injecting Domain Knowledge into DL Models, ML4H December, 2022
- **Reviewer**, Machine Learning for Health (ML4H) September, 2022
- **Reviewer**, Time Series for Health (NeurIPS workshop) September, 2022

LEADERSHIP, VOLUNTEERING, & FUNDRAISING

Department Senator & CMU Representative CMU Graduate Student Association Sept 2021 - Current
Founding director of the CMU x Pitt CompBio Hackathon Pittsburgh, PA

President, Iron Phi, Phi Delta Theta, Washington Alpha Nov 2017 - Nov 2019
Fundraised $49,000 for the Service for Sight Foundation Seattle, WA
Fundraised $1,000 for the Live Like Lou Foundation

Eagle Scout, *Boy Scouts of America* Nov 2015
Construction of the U. Texas Field Laboratory's Research Pier Austin, TX

Run-to-the-Sun 90 Mile Overnight Relay, Beyond Batten Foundation Apr 2015
$500 to the Beyond Batten Foundation Austin, TX

MEDIA (Available on my YouTube, linked at calebellington.com)

- Contextualized Graphical Models Reveal Sample-Specific Transcriptional Networks for 7000 Tumors (CSHL) Nov 2022
- Multi-task Learning of Consistent-Order Causal Graphs (NeurIPS) Dec 2021
- Mixed Feelings: Recognizing, Representing, and Interpolating Human Emotion Dec 2020
- Deep Learning for Contact Prediction in *de novo* Protein Models (UW Research Symposium) May 2020
- Bioscience Python Tutorials 1-4: Protein Sequence Alignment CLI (YouTube) May 2020

SKILLS

Programming: Python (9 yrs), Java (7 yrs), C++ (2 yrs), PyTorch (3 yrs), Tensorflow/Keras (2 yrs), MySQL (2 yrs), React.js (2 yrs), C, Ruby, Django, Git, Superset, AWS, Jira, Linux/Unix, Bash
Languages: Intermediate I Chinese