DANIEL ZEIBERG

Cambridge, MA 02141
zeiberg.d@northeastern.edu ⋄ (609) 410-5967 ⋄ linkedin.com/in/danielzeiberg ⋄ dzeiberg.github.io

RESEARCH INTEREST

Machine Learning
Develop machine learning methods to learn predictive models from datasets with statistical bias
Computational Biology
Discover functional and phenotypic effects of protein coding variants and quantify their association with human diseases

EDUCATION

PhD, Computer Science Expected 05/2024
Northeastern University
Boston, MA 02115

Bachelor of Science in Engineering, Computer Science 2014 - 2018
University of Michigan, Ann Arbor, MI Overall GPA: 3.83/4
Minor in Statistics
Graduated summa cum laude from Engineering Honors College
Member of Eta Kappa Nu and Tau Beta Pi

EXPERIENCE

Graduate Student Researcher 01/2019 - Present
Northeastern University, Advised by Predrag Radivojac *Boston, MA*

- Derived fast non-parametric algorithm to estimate class-prior in positive-unlabeled datasets implemented in Python and Matlab
- Developed novel methods to correct for statistical bias and improve predictive models in semi-supervised classification settings
- Devised high-throughput end-to-end pipelines using protein language models and structure predictors to engineer features for millions of variant calls
- Trained machine learning models to associate protein-coding genetic variants with rare diseases, leading to 17% improvement in classification performance

Graduate Research Assistant 09/2018 - 12/2018
Northeastern University, Advised by Rose Yu *Boston, MA*

- Developed deep-learning-based sequence-to-sequence models using Pytorch and Tensorflow to forecast spatiotemporal data

Undergraduate Research Assistant 05/2017 - 07/2018
University of Michigan, Advised by Jenna Wiens *Ann Arbor, MI*

- Trained a state-of-the-art machine learning model that stratifies hospital patients for their risk of developing Acute Respiratory Distress Syndrome, using electronic health records

Software Defined Core Network Engineer 06/2016 - 08/2016
Comcast *Philadelphia, PA*

- Developed a network health dashboard using Python and Javascript that displays metrics and outages

- Deployed product used widely throughout Comcast's network engineering division
- Automated Comcast's IP address cleanup workflow, managing millions of IPs

Engineering Analysis Intern 05/2015 - 07/2015
Comcast *Philadelphia, PA*

- Created visualizations using D3.js to model On-Demand data flow that were used by business operations personnel to make decisions on network scaling
- Taught myself JavaScript and data visualization best practices

PUBLICATIONS

- Zeiberg, Daniel, Shantanu Jain, and Predrag Radivojac. "Leveraging structure for improved classification of grouped biased data." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 37. No. 9. 2023.
- Zeiberg, Daniel, Shantanu Jain, and Predrag Radivojac. "Fast nonparametric estimation of class proportions in the positive-unlabeled classification setting." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. No. 04. 2020.
- Zeiberg, Daniel, et al. "Machine learning for patient risk stratification for acute respiratory distress syndrome." PloS one 14.3 (2019): e0214465.
- Igvf Consortium. "The Impact of Genomic Variation on Function (IGVF) Consortium." ArXiv (2023).
- Stenton, Sarah L., et al. "Critical assessment of variant prioritization methods for rare disease diagnosis within the Rare Genomes Project." medRxiv (2023): 2023-08.
- Stenton, Sarah L., et al. "Critical assessment of variant prioritization methods for rare disease diagnosis within the Rare Genomes Project." medRxiv (2023): 2023-08.
- Chen, Yile, et al. "Multi-objective prioritization of genes for high-throughput functional assays towards improved clinical variant classification." PACIFIC SYMPOSIUM ON BIOCOMPUTING 2023: Kohala Coast, Hawaii, USA, 37 January 2023. 2022.
- Lugo-Martinez, Jose, et al. "Classification in biological networks with hypergraphlet kernels." Bioinformatics 37.7 (2021): 1000-1007.

AWARDS, PRESENTATIONS, AND REVIEWING

- Most Likely To Have Transformative Scientific Impact - Michigan Institute for Data Science *October 2017*
- Invited Speaker at Michigan Institute for Health Analytics and Medical Prediction *October 2017*
- Reviewer for Intelligent Systems for Molecular Biology *February 2020*

TEACHING EXPERIENCE

Northeastern University 2019-2023
Teaching Assistant *Boston, MA*

- Supervised Machine Learning (Fall 2023)
- Machine Learning (Spring 2020, Fall 2022)
- Data Mining Techniques (Fall 2021)
- Discrete Structures (Spring 2019)

Lavner Camps Summer 2018
Technology Instructor *Cherry Hill, NJ*

- Taught elementary and middle school students the principles of programming, and artificial intelligence