

Nathan LaPierre

NathanL2012@gmail.com

(571) 839-5008

EDUCATION

PhD in Computer Science, 2016 - 2021/2022 (Projected)
University of California, Los Angeles

M.S. in Computer Science, 3.74 GPA, March 2019
University of California, Los Angeles

B.S. in Applied Computer Science, 3.86 GPA (*magna cum laude*), December 2015
George Mason University

PUBLICATIONS

N. LaPierre*, K. Collins*, J. Rotman, and E. Eskin, "Identifying Causal Variants by Fine Mapping Across Multiple Studies," *in preparation*.
(* Authors contributed equally)

N. LaPierre, D. Koslicki, E. Eskin, and S. Mangul, "Metalign: Efficient alignment-based metagenomic profiling via containment min hash," *in preparation*.

N. LaPierre, C. Ju, G. Zhou, and W. Wang, "MetaPheno: A Critical Evaluation of Deep Learning and Machine Learning in Metagenome-Based Disease Prediction," *Methods*, Mar. 2019, In Press.

N. LaPierre, R. Egan, W. Wang, and Z. Wang, "MiniScrub: de novo long read scrubbing using approximate alignment and deep learning," *bioRxiv*, pp. 433573, Oct. 2018.

N. LaPierre, S. Mangul, M. Alser, I. Mandric, N.C. Wu, D. Koslicki, and E. Eskin, "MiCoP: Microbial Community Profiling method capable of detecting low abundance viral and fungal organisms in metagenomic samples," *bioRxiv*, pp. 243188, Jan. 2018. Accepted at BMC Genomics.

M. A. Rahman, **N. LaPierre**, H. Rangwala, and D. Barbara, "Metagenome sequence clustering with hash-based canopies," *Journal of bioinformatics and computational biology*, vol. 15, no. 6, pp.1740006, Oct. 2017.

M. A. Rahman, **N. LaPierre**, and H. Rangwala, "Phenotype Prediction from Metagenomic Data Using Clustering and Assembly with Multiple Instance Learning (CAMIL)," *IEEE/ACM transactions on computational biology and bioinformatics*, Oct. 2017.

N. LaPierre, M. A. Rahman, and H. Rangwala, "CAMIL: Clustering and Assembly with Multiple Instance Learning for Phenotype Prediction," in *IEEE International Conference on Bioinformatics and Biomedicine*, Shenzhen, China, 2016.

N. LaPierre and H. Rangwala, "Predicting Clinical Phenotype using OTU-based Metagenome Representation," in *IEEE International Conference on Data Mining Workshop on Biological Data Mining and its Applications in Healthcare*, Atlantic City, New Jersey, 2015, pp. 156-163.

- PRESENTATIONS**
- “MiniScrub: de novo long read scrubbing using approximate alignment and deep learning” in Amazon Web Services - UCLA Computational Medicine Symposium, February 1, 2019
 - “MiniScrub: de novo long read scrubbing using approximate alignment and deep learning” in American Society for Human Genetics Annual Meeting, October 19, 2018
 - “MiCoP: Microbial Community Profiling method capable of detecting low abundance viral and fungal organisms in metagenomic samples” in American Society for Human Genetics Annual Meeting, October 20, 2017
 - “CAMIL: Clustering and Assembly with Multiple Instance Learning for Phenotype Prediction” in IEEE International Conference on Bioinformatics and Biomedicine, December 16, 2016
 - Won a conference travel grant sponsored by NSF
 - “Predicting Clinical Phenotype using OTU-based Metagenome Representation” in IEEE International Conference on Data Mining workshop on Biological Data Mining and its Applications in Healthcare, November 14, 2015
 - Won a travel grant from the Undergraduate Student Travel Fund of the Office of Student Scholarship, Creative Activities, and Research at GMU
 - “Developing a Computational Pipeline for Metagenomic State Classification with Feature Engineering” in Volgenau School of Engineering Undergraduate Research Celebration, April 2015
 - Won Outstanding Undergraduate Research Project Award for poster

RESEARCH EXPERIENCE

- PhD Student*
University of California, Los Angeles September 2016 - Present
- Working with Professor Wei Wang and Professor Eleazar Eskin
 - Developing method for identifying causal variants (statistical fine mapping) by leveraging information from multiple studies using a Bayesian approach
 - Developed a method, that uses deep learning to improve long sequencing read quality, in collaboration with scientists from the Joint Genome Institute
 - Developed a method, accepted at BMC Genomics, that performs accurate abundance profiling of viruses and fungi in metagenomic samples
- Graduate Research Assistant (GRA) / Student Researcher*
George Mason University January - August 2016 (June-August as GRA)
- Worked with Professor Huzefa Rangwala
 - Developed CAMIL, a pipeline that uses multiple instance learning techniques based on whole metagenome shotgun sequence reads to predict whether or not a patient has a disease.
 - CAMIL paper accepted into IEEE BIBM 2016 (19% acceptance rate).
 - Second author of paper on using canopy clustering and locality sensitive hashing to reduce clustering time for biological datasets.
- Predicting Clinical Phenotype using OTU-based Metagenome Representation*
George Mason University January 2015 - November 2015
- Worked with Professor Huzefa Rangwala
 - Developed a computational pipeline that uses clustering and classification methods to quickly and accurately predict whether a patient has a disease based on a case/control metagenomic dataset

- Paper accepted for publication, and poster presentation won Outstanding Undergraduate Project Award (see publications and presentations sections)

WORK EXPERIENCE

Graduate Student Researcher Summer 2017
DOE Joint Genome Institute / Lawrence Berkeley National Lab

- Developed a method that uses a Convolutional Neural Network (Deep learning method) to improve Oxford Nanopore long read quality *de novo*
- Method improves both read accuracy and downstream *de novo* assembly
- Helped pioneer use of deep learning within the lab

Security Engineering Intern Summers 2013-2015
Sony Corporation of America

- Wrote secure and scalable software and worked with Big Data in order to help analyze, detect, and prevent attacks on Sony's networks
- Used Python, Javascript (Node, Express, Meteor), MongoDB, X/HTML, CSS
- Developed a workplace communications system using Meteor.js; solo project
- Assisted in the development of a web application based on a searchable database system using Node.js, Express.js, and MongoDB
- Assisted in the development of a network forensics system; wrote backend python scripts and XML web layouts

TEACHING EXPERIENCE

Undergraduate Teaching Assistant Fall 2014 - Spring 2015
CS 306 - Computer Law and Ethics, Computer Science Department, George Mason University

- Received Outstanding Undergraduate Teaching Assistant Award
- Assisted students with legal research, writing, and oral communication
- Responsible for grading student assignments worth 25% of their grade; one of the only Undergraduate Teaching Assistants entrusted with this responsibility

AWARDS, FELLOWSHIPS, SCHOLARSHIPS

Selected Honors

- *Honorable Mention in NSF Graduate Research Fellowship*, 2015-16
- *Outstanding Undergraduate Student Award*, given to overall best undergraduate student in Computer Science at George Mason University, May 2016
- *Outstanding Academic Achievement Award*, given for outstanding performance in Computer Science, May 2016
- *Outstanding Undergraduate Research Project* for presentation of my research project at Volgenau School of Engineering Undergraduate Research Symposium, April 2015
- *Outstanding Undergraduate Teaching Assistant* for two semesters of excellence as a teaching assistant, April 2015
- *Dean's List* every semester

Fellowships and Merit-Based Scholarships

- *ModEling and uNdersTanding human behaviOR (MENTOR) NSF Training Grant*, 2018-19 academic year at UCLA
- *NIH T32 Doctoral Training Fellowship*, 2016-18 academic years at UCLA
- *Bersoff Endowed Scholarship*, Outstanding Academic Achievement, Awarded twice in 2015 and 2014
- *SWIFT Scholarship*, Outstanding Academic Achievement, Awarded in 2015

Honors Societies and Organizations

- *Honors College* at George Mason University
- *Alpha Lambda Delta Honor Society*

- *Golden Key International Honour Society*

Technical Competitions

- *Top 5 at VTHacks*, a software development competition at Virginia Tech with over 45 teams, April 2014
- *2nd Place in the Technical Innovation Challenge*, a week-long competition at George Mason University to design a viable software product, jointly refereed by Computer Science and Business Departments, October 2014

VOLUNTEER EXPERIENCE

- Executive Curriculum Planner and Mentor, Community Programming Initiative
George Mason University Honors College and SRCT 2014 - 2015
- Volunteer effort to teach basic programming to local 5th-8th grade students
 - Planned and developed parts of the curriculum for the sessions, such as designing games and hints to help the students create them
 - Mentored the elementary and middle school students during the sessions

COMPUTATIONAL SKILLS

Languages / Scripting: Python, R, C, Matlab/Octave, bash

Deep Learning Frameworks: Keras, TensorFlow, PyTorch

Other Technologies: Git, Docker, LaTeX

Operating Systems: Linux, Windows, macOS

Experience with:

- Machine Learning Algorithms: Deep Learning, Linear Regression, Logistic Regression, SVMs, Random Forests, etc
- Machine Learning Applications: Clustering, Classification, Bioinformatics
- Bioinformatics methods experience: statistical fine mapping, alignment, assembly, read error correction, metagenomics, abundance profiling
- SGE and Slurm grid/cluster computing systems

MEMBERSHIPS AND LEADERSHIP

- Student-Run Computing and Technology (SRCT)* 2013 - 2016
- Student organization at George Mason University that works on software projects and competitions to benefit the university and broader local community
 - Secretary and Member of Executive Board, Fall 2015 Semester