

JONATHAN DEATON

1125 Park Place Apt 215 San Mateo, CA 94403 • jdeaton@google.com • (415) 531-1568 • [linkedin.com/in/deatonjonathan](https://www.linkedin.com/in/deatonjonathan)

EDUCATION

- 01/18 – 06/19 **Stanford University**, Stanford CA – MS in Computer Science, Artificial Intelligence
- Cumulative GPA – 3.957 / 4.0
 - Machine Learning, Reinforcement Learning, Algorithmic Analysis, Operating Systems
- 09/12 – 03/18 **Stanford University**, Stanford CA – BS in Bioengineering with Honors
- Cumulative GPA – 3.839 / 4.0, Departmental Honors for GPA, honors research thesis
 - Tau Beta Pi National Engineering Honors Society member (top 20% of engineering class)
-

EXPERIENCE

- 09/19 – Present **Software Engineer**, Google Health, Dermatology
- Productionized ML computer-vision [dermatological-condition classification system](#), gaining regulatory approval as CE marked Class I medical device in the EU
 - o Solved technical modeling hurdles regarding missing data and healthy skin recognition
 - o Lead development of software systems for evaluation and validation of ML models
 - Wrangled difficult medical datasets
 - Designed and conducted research studies improving and advancing understanding of state-of-the-art ML models, publishing research findings.
 - Hosted research intern in project on negative interference in multi-dataset domain generalization
- 09/18 – 06/19 **Course Assistant**, Department of Computer Science, Stanford University
- CS 161: Design and Analysis of Algorithms (Fall Quarter)
 - CS 140: Operating Systems (Spring Quarter)
- 06/18 – 09/18 **Software Engineering Intern**, Bigstream Solutions Inc., Mountain View, CA
- Identified and presented technical and business opportunities in big-data genomics
 - Established genomic data pipelines for performance benchmarking Bigstream technology
 - Accelerated genomic sequencing duplicate-read marking algorithm using Apache Spark SQL
- 01/18 – 04/18 **Graduate Research Assistant**, Stanford Network Analysis Platform (SNAP), Stanford University
- [Transformed 2TB of Reddit user data](#) using Python, Pandas, multiprocessing and Redis
 - Enabled development of models for predicting Reddit user interactions
- 06/17 – 12/17 **Software Test Engineer**, Digital Systems, Illumina Inc., San Diego, CA
- Developed automated testing tools in C# to support continuous integration processes
 - Reduced automated testing code size by 40% (~3,000 lines) while adding 25% (11) more tests
 - Diagnosed 5 software bugs, including a major security vulnerability in release candidate software
 - Developed web application for tracking continuous integration metrics
- 06/14 – 06/17 **Undergraduate Bioengineering Researcher**, Quake Lab, Stanford University
- Developed experimental analysis and visualization tools in Python and MATLAB
 - Conceived of, performed, and analyzed molecular biology / high-throughput sequencing experiments
 - Designed and fabricated experimental devices, procedures, and software
 - Communicated scientific findings in academic conferences, lab meetings, and publications
- 06/16 – 09/16 **Research Associate Intern**, Protein Engineering Group, Illumina Inc., San Diego, CA
- Developed novel algorithm for quantifying decay in DNA sequencing run data quality in Python
 - Experimented with optimizing assays for human exome sequencing
 - Automated experimental data analysis with Python
- 09/12 – 04/16 **Division I Varsity Athlete**, Men's Gymnastics, Stanford University
- Balanced rigorous course load with 25+ hours of training per week
 - Performed well in high pressure situations
 - 2013-15 First Team NCAA Academic All American
-

PUBLICATIONS

- Big Self-Supervised Models Advance Medical Image Classification. Azizi S, Mustafa B, Ryan F, Beaver Z, Freyberg F, **Deaton J**, Loh A, Karthikesalingam A, Kornblith S, Chen T, Natarajan V, Norouzi M. International Conference on Computer Vision 2021. [arXiv](#)
- Addressing the Real-world Class Imbalance Problem in Dermatology. Weng WH, **Deaton J**, Natarajan V, Elsayed GF, Liu Y. NeurIPS Machine Learning for Health Workshop 2020. [arXiv](#)
- Underspecification Presents Challenges for Credibility in Modern Machine Learning. D'Amour A, Heller K, et al. [preprint arXiv](#)
- Mini-Metagenomics and Nucleotide Composition Aid the Identification and Host Association of Novel Bacteriophage Sequences. **Deaton J**, Yu FB, Quake S. Advanced Biosystems 2019. [Journal paper](#)

PROJECTS

- 01/19 – 09/19 **AgarLE**, High-performance Multi-Agent Reinforcement Learning Environment
- Implemented and wrapped performance-oriented C++ game engine in OpenAI gym interface for RL algorithm development in Python (<https://github.com/jondeaton/AgarLE>)
 - Implemented in JAX modern deep-RL algorithms including PPO, A2C, DQN, GAE
- 03/18 – 06/18 **Brain Tumor Segmentation using CNNs**, Deep Learning Final Project, Stanford
- Designed deep neural network for segmentation of brain tumors from 3D MRI scans
 - Implemented and trained [fully-convolutional UNet architecture using TensorFlow](#)
 - Demonstrated near state-of-the-art performance with ~0.9 dice coefficient on validation set
- 10/15 – 03/17 **Bacteriophage Classification Algorithm**, Undergraduate Honors Thesis
- Designed and implemented [machine learning algorithm in Python to detect viral DNA sequences](#)
 - Discovered and characterized over 100 novel viruses
 - Presented at the Physical Biology of the Cell conference in Kona, Hawaii
 - Publication: Deaton, Jonathan, et al. “PhaMers Identifies Novel Bacteriophage Sequences from Thermophilic Hot Springs.” *BioRxiv*, 28 July 2017. [bioRxiv Publication](#)
- 10/16 – 03/17 **Heat Stroke Risk Monitor**, Bioengineering Senior Capstone, Stanford
- Led team of three in designing, building, and testing wearable heat stroke risk monitor
 - Developed and tested [machine learning algorithms in Python](#) to predict user risk in real time
 - Presented at 2017 Tau Beta Pi Engineering Showcase, Rice 360 Global Health Competition Nomination
- Personal Projects**
- [Lisp interpreter in C](#) featuring closures, currying, recursion, deterministic memory management
 - Multi-threaded, [distributed batch processing framework](#) in C++ using OpenMPI
 - Assembled and configured [distributed computing cluster](#) using Arch Linux
-

TECHNOLOGY

- Significant experience/familiarity with Python, C++, C, Linux, JAX, TensorFlow, computer vision, supervised learning, genomics
- Experience with reinforcement learning, PyTorch, high-throughput genomic sequencing, molecular biology