Jonathan Hayase

jhayase@cs.washington.edu | github/PythonNut | (714) 492-0500 | As of October 3, 2023

Education

4th year Ph.D. student at the Paul G. Allen School of Computer Science & Engineering.

B.S., Joint Major in Computer Science and Mathematics from Harvey Mudd College 10/2016 — 05/2020.

Selected Papers

Label Poisoning is All You Need *accepted to NeurIPS 2023*

- Rishi Jha, **Jonathan Hayase**, Sewoong Oh
- We show that label poisoning alone is able to construct backdoor attacks for image classification models with arbitrary image-space triggers.

DataComp: In search of the next generation of multimodal datasets *accepted (oral) to NeurIPS 2023*

- S. Y. Gadre, G. Ilharco, A. Fang, **J. Hayase**, G. Smyrnis, T. Nguyen, R. Marten, M. Wortsman, D. Ghosh, J. Zhang, E. Orgad, R. Entezari, G. Daras, S. Pratt, V. Ramanujan, Y. Bitton, K. Marathe, S. Mussmann, R. Vencu, M. Cherti, R. Krishna, P. W. Koh, O. Saukh, A. Ratner, S. Song, H. Hajishirzi, A. Farhadi, R. Beaumont, S. Oh, A. Dimakis, J. Jitsev, Y. Carmon, V. Shankar, L. Schmidt
- We introduce a comprehensive testbed for multimodal dataset curation and use it to construct DataComp-1B, a dataset which trains CLIP ViT-L/14 to 79.2% zero-shot on ImageNet, beating OpenAI's CLIP ViT-L/14 by 3.7 pp while using the same training procedure and compute.

Few-shot Backdoor Attacks via Neural Tangent Kernels ICLR 2023

- **Jonathan Hayase**, Sewoong Oh
- We use the Neural Tangent Kernel to design backdoor attacks against neural networks using dramatically fewer poisoned examples.

Git Re-Basin: Merging Models modulo Permutation Symmetries oral @ ICLR 2023

- Samuel K. Ainsworth, **Jonathan Hayase**, Siddhartha Srinivasa
- We show that the hidden units of independently trained models can be permuted such that there is no loss barrier between the models in weight space.

Zonotope Domains for Lagrangian Neural Network Verification NeurIPS 2022

- Matt Jordan, **Jonathan Hayase**, Alexandros G Dimakis, Sewoong Oh
- We give tighter bounds for NN verification using Zonotope abstract domains to approximate the dual.

SPECTRE: Defending Against Backdoor Attacks Using Robust Statistics ICML 2021

- **Jonathan Hayase**, Weihao Kong, Raghav Somani, Sewoong Oh
- We defend against backdoor attacks using high dimensional robust mean and covariance estimators.

Patents

Security threat monitoring for a storage system, US10970395B1 2021

- A. Bansal, O. Watkins, **J. Hayase**, N. Bhargava, C. Golden, S. Zhuravlev
- System to detect security threats by analyzing storage access patterns using machine learning.

Skills

Languages: Python, Julia, C, C++, JavaScript, Emacs Lisp, LaTeX
Machine Learning: JAX, PyTorch, FluxML, scikit-learn
Tools: Git, MIP solvers, SAT solvers, Z3, React

Work Experience

Software Engineer, Scotts Miracle-Gro Company, remote 2020

- Created Google Cloud microservices for geolocation, address normalization, SMS, email, job scheduling.
- Created REST API test and documentation repository and microservice starter template.

Software Engineering Intern, Pure Storage, Inc., Mountain View, CA 2018–2019

- Ported Purity Operating Environment to Microsoft Azure.
- Worked on scripts to deploy and manage Azure components using Python.
- Wrote cloud deployment scripts using the Azure Resource Manager and Terraform.

Data Science Intern, UnifyID, San Francisco, CA 2018

- Wrote machine learning models in Python to classify user behavior via cellphone accelerometers.
- Performed exploratory data analysis on several biometric datasets using Julia.

Software Engineering Intern, NovaWurks, Inc., Los Alamitos CA 2017

- Developed a robust, high-performance communication framework for use on satellites in C.
- Operated the hardware integration and mission simulator test bench for the eXCITe DARPA mission, which flew Dec 2018.

Computer Science/Engineering Intern, McKinley Equipment, Anaheim CA 2014–2016

- Proposed and implemented scalable server configuration management and automation.
- Worked on embedded C++ on ARM microprocessors for Internet of Things devices.
- Wrote a network abstraction library for LoRa radios, for use under extreme power draw constraints.

Teaching Experience

Grader and Tutor, Harvey Mudd College 2018-2019

- Tutored other students and graded assignments for Computability & Logic, Advanced Topics in Algorithms, and Mathematics of Big Data

Coursework

Machine Learning: Machine Learning, Deep Learning, Deep Learning Theory, Interactive Learning, Math of Data Science, Advanced Big Data Analysis

Computer Science: Data Structures & Program Development, Programming Languages, Computability & Logic, Scientific Computing, Digital Electronics & Computer Engineering, Advanced Topics in Algorithms, Random Algorithms

Mathematics: Positive Definite Matrices, Optimal Transport, Seminar in Differential Geometry, Advanced Linear Algebra, Measure Theory, Representation Theory, Knot Theory

Teaching Assistant, Harvey Mudd College 2018-2019

- Served as a teaching assistant for Seminar in Differential Geometry and Advanced Linear Algebra.

Honors & Awards

- National Science Foundation Graduate Research Fellowship Program (2021–2026)
- Interdisciplinary Contest in Modeling, Meritorious Winner (2019)
- Pure Storage Hackathon Grand Prize (2018)
- 5C Hackathon, Best Game (2017)
- MuddHacks, Top Six Teams (2016)
- 5C Hackathon Intermediate Division, 1st Place (2016)
- Harvey S. Mudd Merit Scholarship (2016–20)
- Harvey Mudd College Dean’s List (2017–present)