Huey Sun

+1 (408) 242-9951 | huey3sn@gmail.com | github.com/itshuey

Education

University College London — London, U.K.
MSc Computational Statistics and Machine Learning — *2023 – 2024*

- Final Mark: 77.7% **(Distinction)**.
- Thesis: Thinking Forwards, Backwards, and in Code (83%), advised by Prof. Pontus Stenetorp and Dr. Yihong Chen. Developed a taxonomy of failure modes in multi-step tool use, introduced backwards and code-structured reasoning traces, fine-tuned a model to **state-of-the-art performance at 7B** on ToolSandbox.
- Selected Coursework: Bayesian Deep Learning, Applied Deep Learning, Machine Vision, Probabilistic & Unsupervised Learning (Gatsby Unit), Supervised Learning, Statistical NLP

Pomona College — Claremont, CA
BA Mathematics, Minor in Computer Science — *2016 – 2020*

- Graduated cum laude; GPA: 3.85 overall, **3.97 major/minor**.
- Thesis: The Limits of Transfer Learning, advised by Prof. Ghassan Sarkis.
- Honors: **Llewellyn Bixby Mathematics Prize**; 5× Pomona Scholar (top 25%).
- Study abroad: Budapest Semesters in Mathematics (Spring 2019)
- Selected Coursework: Machine Learning, Computational Statistics, Combinatorial Optimization, Probability, Real Analysis, Theory of Computation, Advanced Programming, Algorithms.

Research Experience

Cohere Labs — San Francisco, CA
Research Scholar — *2025*

- Researched **personalization of large language models** by designing embedding-based representations of preference data. Identified latent preference patterns across users and trained lightweight, mode-specialized adapters that improved personalization performance at reduced cost.
- Implemented new architectures, loss functions, and training pipelines in PyTorch and JAX, built tooling for evaluation and deployment workflows, **fine-tuned models with RLHF** training.
- Curated 100+ datasets and rebuilt our API-data pipeline (querying → cleaning → prep). Contributed to core libraries for synthetic data generation, inference, and adapter fusion.
- *Ongoing work:* manuscript in preparation for 2026.

Harvey Mudd College, Dept. of Computer Science — Claremont, CA
Clinic Research Assistant — *2019 – 2020*

- Worked on theoretical foundations of transfer learning through the lens of algorithmic search. Proved bounds using information-theoretic tools under advisor Prof. George Montañez.
- Co-authored two peer-reviewed publications.

Pomona College, Dept. of Mathematics — Claremont, CA
Research Assistant (Erica Flapan) — *Summer 2018*

- Studied spatial graphs and their applications to molecular structures; developed a taxonomy of synthetic organometallic knot/link structures.

Pomona College, Dept. of Chemistry Claremont, CA
Research Assistant (EJ Crane) *Summer 2017*

- Investigated microbial sulfur reduction via cyclic voltammetry; built custom electrodes and analyzed formation of transient sulfur compounds.

Publications & Preprints

- **H. Sun**, S. Singh, A. Ahmadian, J. A. Campos, B. Ermis, S. Hooker. *Disentangling Preference Soups With Parameter-Efficient Personalization.* Manuscript in preparation (2026).
- **H. Sun**, A. Yong, L. Gilly, F. Jin. *Enhancing Instruction-Following Capabilities in Seq2Seq Models: DoLA Adaptations for T5.* arXiv preprint (2025).
- O. Bridge, **H. Sun**, B. Branyicskai-Nagy, C. D'Ornano, S. Basu. *Diminishing Returns in Self-Supervised Learning.* arXiv preprint (2025).
- J. Williams, A. Tadesse, T. Sam, **H. Sun**, G. Montañez. *Limits of Transfer Learning.* LOD (2020). *(Student authors contributed equally).*
- T. Sam, J. Williams, A. Tadesse, **H. Sun**, G. Montañez. *Decomposable Probability-of-Success Metrics in Algorithmic Search.* ICAART (2020). *(Student authors contributed equally).*

Professional Experience

Microsoft Redmond, WA
Software Engineer II, Cloud+AI *2020 – 2023*

- Designed and scaled cloud microservices for cybersecurity analytics and detections.
- Implemented models for blast radius, performance regression detection, threat vector modeling, and core caching/encryption libraries for Azure services.
- Built a shared compute platform reducing Spark job latency from 5 minutes to seconds, improving live endpoint monitoring throughput.

Wove Technologies San Francisco, CA
Full-Stack Engineering Intern *2019*

- Implemented online multi-arm bandit algorithms and built messaging/analytics features across the stack (React, Ruby, Postgres, Java).

Teaching Experience

Teaching Assistant, Pomona College: Advanced Programming (2019–20), Combinatorics (2018), Organic Chemistry (2018), Genetics (2017). Led office hours, discussion/lab sessions, and grading.

Entrepreneurship

Theatr: Co-founded a marketplace for secondhand theater tickets (20k MAU; millions in annual sales), featured in Forbes, NPR, and Vulture.