Bruce N. Hirwa

Cupertino, CA | bruce.nshutihirwa@gmail.com | linkedin.com/in/brucenh | github.com/brucenh

Education

University of California, Santa Cruz — Santa Cruz, CA
B.S. Computer Science — Fall 2026 — *Expected Jun. 2028*
- Coursework: Data Structures & Algorithms, Discrete Math, Linear Algebra, Probability, Object-Oriented Design

De Anza College — Cupertino, CA
A.S.T. Computer Science & Mathematics — *Sep. 2024 – 2026*

Research & Publications

Incoming MTS — Aug. 2026
LexAxiom
- Working on building the future for business in law

Machine Learning Researcher — Jul. 2025 – May 2026
Algoverse AI Research
- Built hidden-state extraction pipeline on Llama-3.1-8B over 44K+ GPT-4o-labeled poker samples; achieved AUROC ∼0.95 separating bluff vs. honest reasoning traces.
- Engineered SAE v3 (z-score norm + tied decoder) to resolve feature collapse; identified deception circuits via linear probes and bidirectional activation patching (100% patch success rate).
- Co-authoring paper submitting to **Mechanistic Interpretability Workshop**; extending to sandbagging detection and Llama-3.1-70B.

Google x TechWise Fellow — Mar. 2024 – May 2025
TalentSprint (Google & Carnegie Mellon University)
- Selected 1 of 124 nationwide for competitive fellowship; completed intensive training in Python, DSA, full-stack development, and ML fundamentals with production-level project deliverables.

Projects

JurisPrudence - Pro Se Legal Assistant | *Python, Claude API, RAG, Streamlit*
- Built an LLM-driven assistant that drafts and explains small-claims / unlawful-detainer responses for self-represented litigants, grounding generations in [statute/rule set, e.g. CA Code of Civil Procedure] via retrieval.
- Designed a citation-verification eval that flags hallucinated or non-existent authorities before output, cutting fabricated citations from 15% to sub 4% on a 30-case test set.
- Built to close the access-to-justice gap for litigants who cannot afford counsel.

Perceptron from Scratch | *Python, NumPy*
- Implemented single-layer neural network without high-level ML libraries.
- Built forward propagation, loss computation, gradient descent, and evaluation pipeline from first principles.
- Benchmarked performance against classical classifiers.

Leadership & Activities

VP, Competitive Programming Club (ICPC), De Anza College — Jan. 2025 – Mar. 2026

Honorable mentions at **ICPC PacNW 2025**. Lead weekly algorithm sessions for 20+ members targeting Regionals

President, Black Student Union, De Anza College | **TA** Statistics 10 | **VP**, Phi Theta Kappa HIA

Interests: Hackathons (CalHacks, UC Davis Hacks, De Anza Hacks (5 total)), Coke Zero

Technical Skills

Languages: Python, Java, C++ **Agentic Tools:** Claude Code, Cursor, prompt engineering

AI/ML: PyTorch, NumPy, pandas, scikit-learn, SBERT, SAEs, linear probes, LLM APIs (Claude, GPT-4o)

Systems & Web: Git, Linux, Docker, REST APIs **Domains:** OOP/Design Patterns, Mechanistic Interpretability, Algorithms & Data Structures