

Michael Lavery

Berkeley, CA | P: 1 (669) 252 - 6813 | mlavery@berkeley.edu

EDUCATION

University of California, Berkeley

Electrical Engineering and Computer Science

Coursework: Operating Systems, AI, Machine Learning, Random Processes, LLM Agents

Berkeley, CA

Dec 2025

GPA: 3.9 / 4

EXPERIENCE

Improving Robustness in Machine Learning Models, Researcher

Jun 2025 – Present

- Lead a research project that produced protections against prompt injection that generalize to agentic settings
- Addressed AI safety limitations by creating a dataset targeting safety specifically for reasoning models
- Determined that applying the GRPO method to Qwen 3 models results in a strong defense while preventing regressions

Berkeley AI Research Lab (BAIR), Student Researcher

Jun 2023 – Jun 2025

- Improved LLM instruction following by 60% in David Wagner's group using our synthetic SFT and DPO dataset
- Increased experiment variety by 2x through adding OLMo model architecture to our in-house LLM framework
- Determined the best approach to evaluating instruction following by organizing groups of common LLM failure modes
- Ensured fine-tuned models maintain utility by contributing model capability evaluation pipelines
- Paper submissions: NeurIPS 2024 workshop

Berkeley AI Research Lab (BAIR), Machine Learning Intern

Jun 2024 – Aug 2024

- Hired by Trevor Darrell's group to fulfill a satellite image detection models contract
- Pretrained experimental detectors to 0.18 mAP (SOTA number) by training on xView1 aerial data
- Provided library modifications with documentation to support our experimental detectors
- Successfully concluded the project by delivering Docker containers for training, finetuning, and inference to the contractee

Lawrence Berkeley National Lab (LBNL), Intern

Jun 2023 - Jul 2023

- Developed a Python FastAPI file download service to provide automatic software updates for new systems

PROJECTS

Infinite Range Delivery Robot - <https://github.com/thavens/mobot>

- Built a 4G-enabled Raspberry Pi hoverboard robot to do my bidding remotely
- Produced a custom fault tolerant TURN server on AWS (Amazon Web Services) to relay data from control client to robot
- Streamed ultra-low latency video over 4G with C++, leveraging Libavcodec's efficient video compression

DQN 2048/Cartpole - <https://github.com/thavens/Gym-RL.git>

- Created a project to teach a reinforcement learning series at the Developers' Guild
- Taught students from the ground up by developing a Deep Q-Network implementation with PyTorch from scratch

ACTIVITIES

CoDirector, De Anza Hacks

Cupertino, CA

- Led a team of 15 to organize De Anza's first 2-day in-person hackathon for 100+ attendees

Vice President, Developers' Guild

Cupertino, CA

- Taught 100 students RNN, CNN, DQN, DDQN, and GAN
- Created course content like autograd graph visualization in Micrograd and PyTorch Google Colab training notebooks

SKILLS

Languages: Python, Java, C++, Rust

Technologies: PyTorch, Tensorflow, Hugging Face, Transformers, Accelerate, FastAPI, Git/Github, Docker, Linux, Slurm