Jay Yeung

650-293-7697 | jayyeung@berkeley.edu | linkedin.com/in/jayyeung | https://github.com/JayYeung | jayyeung.vercel.app

EDUCATION

University of California, Berkeley — May 2027

Bachelor's Degree Computer Science, Double in Data Science — *GPA: 4.0*

- **Relevant Coursework:** Probability and Random Processes (EECS 126), Operating Systems (CS 162), Efficient Algorithms (CS170), Advanced Matrix Computations (M221)
- **Relevant Teaching:** Discrete Mathematics and Probability Theory (CS70) Teaching Assistant (UGSI 2)
- **Awards:** AIME Qualifier (2021), LeetCode Guardian (2300 rating, top 0.4%) my profile

EXPERIENCE

Meta — Menlo Park, CA

Software Engineer Intern — *May 2025 – Aug 2025*

- Model efficiency and optimization analysis for Messenger ML Infrastructure Team

Berkeley AI Research (BAIR), LLM Knowledge Editing and Localization — Berkeley, CA

Undergraduate Researcher at Berkeley Speech Group — *Jan 2025 – May 2025*

- Casual Tracing and Tuned Lens to localize of factual recall and Parts of Speech encoding
- Used Paramater efficient finetuning (PEFT) and Low Rank Adaptation (LoRA) to finetune models.
- Evaluated using Siamese Models and Reinforcement Learning Human Feedback (RLHF) against baseline models.

Cadence Design Systems — San Jose, CA

Hardware Machine Learning Intern — *May 2024 – Aug 2024*

- Developed Hypergraph Neural Network (HGNN) and Graph Attention Network (GAT) for analog model classification
- Optimized simulations for Spectre (EDA tool) under the mixed analog-digital signal (AMS) team

Walmart — Berkeley, CA

Software Engineer — *Jan 2024 – May 2024*

- Object detection (YOLOv8), tracking (DeepSORT), permanence and crowd counting for Walmart inventory tracking.
- Object classification using Siamese Models and Detectron2 for 50+ classes of inventory items.

GitHub — Berkeley, CA

Software Engineer — *Sep 2023 – Jan 2024*

- Automated classification of 40,000 repositories with relevant topics via NLP and LLMs, saving 1000+ hours
- Achieved 96.5% using Multi-Modal Neural Network Class Classification accuracy

NXP Semiconductors N.V. — San Jose, CA

Hardware Engineering Intern — *Jun 2021 – Jun 2022*

- Used Kernighan–Lin algorithm to speed up the run time from 50 to 4500 gates in seven hours

PROJECTS

TreeHacks 2025

- Won Tesla's Autonomy Track Grand Prize (1st place), and Tesla's Excellence Prize (3rd place). $\sim$ 8k in prizes.

Datathon 5.0 C-Light at Berkeley

- Won a Berkeley Datathon using Generative AI for eye retinal scans
- Used GANS, Auto-encoders, Latent Space Sampling, and Embedding Interpolation to generate synthetic videos

Snap Detection

- Built a neural network that detects finger snaps that can be used to activate household appliances
- Hosted the project on GitHub and created a YouTube demo

TECHNICAL SKILLS

Programming Languages: Python, JavaScript, HTML, CSS, Java, C++, Perl, TypeScript

Frameworks: TensorFlow, React, Next.js, OpenGL, Flask, Django, Bootstrap, Express.js

Libraries: Git, NumPy, Pandas, SciPy, XGBoost, Matplotlib, PyTorch, Sk-Learn, Scrapy, LangChain, OpenCV, BeautifulSoup, MediaPipe, PyAutoGUI, pickle, tabula, NLTK, RegEx, pyimgui, seaborn