

Jay Yeung

650-293-7697 | jayyeung@berkeley.edu | [linkedin.com/in/jayyeung](https://www.linkedin.com/in/jayyeung) | <https://github.com/JayYeung> | jayyeung.vercel.app

EDUCATION

University of California, Berkeley

Expected May 2026

Bachelor's Degree Computer Science

GPA: 4.0

- **Relevant Coursework:** Combinatorial Algorithms and Data Structures (CS270), Efficient Algorithms and Intractable Problems (CS170), Advanced Matrix Computations (M221)
- **Relevant Teaching:** Discrete Mathematics and Probability Theory (CS70) Tutor, Structure and Interpretation of Computer Programs (CS61a) Academic Intern

EXPERIENCE

Cadence Design Systems

San Jose, CA

Hardware Machine Learning Intern

May 2024 – Present

- 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 Intern

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.

Berkeley AI Research (BAIR), LLM Conversational Style Cloning

Berkeley, CA

Undergraduate Researcher

Jan 2024 – May 2024

- Pioneered state-of-the-art transformers with over 7B parameters, specializing in mimicking diverse speech styles.
- 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 RAG models.

GitHub

Berkeley, CA

Software Engineering Intern

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

Broker Brain, LLC

Park City, UT

Co-Founder

Mar 2023 – Nov 2023

- Enhanced house search experience through ChatGPT 4, RAG, and DALL-E 3 on askmarcie.com
- Automated processing over one million steel listings using OCR and GPT-based techniques

NXP Semiconductors N.V.

San Jose, CA

Hardware Engineering Intern

Jun 2021 – Jun 2022

- Leveraged PyTorch to drastically reduce the ATE debug time required in creating test-patterns
- Used Kernighan–Lin algorithm to speed up the run time from 50 to 4500 gates in seven hours

PROJECTS

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](https://github.com) and created a [YouTube demo](https://www.youtube.com/watch?v=...)

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