

CHRISTOPHER MICHAEL VU

562-533-4524 | christopher.vu@berkeley.edu | [linkedin.com/in/chrismichaelvu](https://www.linkedin.com/in/chrismichaelvu)

EDUCATION

University of California, Berkeley

B.A. Computer Science, Data Science | Dean's List

Berkeley, CA

Grad: May 2028

- **GPA:** 4.0

- **Coursework:** Linear Algebra, Discrete Math/Probability, Computer Programming, Data Science 1 & 2, NLP

EXPERIENCE

Basis

San Francisco, CA

Machine Learning Engineer Intern

July 2026 – Present

- Trained a text-audio dual encoder (WavLM+RoBERTa, contrastive fine tune) to embed vocal emotion for dataset tasks (labeling, quality rating), beating SOTA audio embeddings by 24% on our benchmark
- Ran LoRA ablations (synthetic data, self-distillation) to prune unnecessary components and finalize the architecture, cutting training cost by ~31%, then benchmarked existing audio encoders on a 400h corpus

AMD

Berkeley, CA

Data Scientist (Contract through Data Science Society at Berkeley)

February 2026 – May 2026

- Built a GPU trace analysis library that matches kernel patterns across vLLM builds and generates diagnostic plots/stats, becoming the standard trace ingestion method for 8 team members and processing 12M+ kernels
- Trained a 1D conv. autoencoder to flag regressions, lifting recall from 55% to 92% on a synthetic benchmark

UC Berkeley School of Information

Berkeley, CA

Undergraduate Researcher

February 2026 – Present

- Built a pairwise LLM-as-judge pipeline to score local laws on interpretive axes (e.g. severity) via TrueSkill ratings, then distilled 160k+ comparisons into a BERT regressor for automated scoring across 1.7M+ statutes
- Built a Next.js/Postgres annotation platform to human-validate the pipeline, collecting 2k+ legal judgements

Hallmate

Cerritos, CA

Fullstack Developer (Freelance)

June 2024 – May 2025

- Developed a Django/Nuxt digital hall pass, hosting 400+ students and reducing hall time by ~30%
- Integrated a multi-source RAG pipeline with LLM-generated reports, procedurally calculated statistics, and stringified database entries to enable natural language database interaction for teachers

SimplyCS

Cerritos, CA

Founder & President

September 2023 – Aug 2025

- Directed a cabinet of 8 to teach beginner-friendly CS curricula to 100+ grade/middle school students
- Secured \$10,500+ and partnered with C2KSD to provide refurbished laptops to low-income students
- Boosted revenue 25% through direct outreach and social media campaigns, reaching 2,500+ families

PROJECTS

CaseLink

Berkeley, CA

Berkeley Data Science Symposium 2025

December 2025

- Presented a legal case text similarity-search engine at UC Berkeley Data Science Symposium
- Built a FAISS search index for ~25k legal docs embedded using a domain-fine-tuned SBERT model
- Addressed explainability via saliency maps generated with sliding token-window occlusion analysis

SafeRoute

Berkeley, CA

Cal Hacks 2025

October 2025

- Built a real time crime map, combining both user-reported incidents and live scraping pipelines
- Implemented a routing algorithm optimized for both safety and travel time using crime density data

SKILLS AND INTERESTS

ML/Data: Python, PyTorch, TensorFlow, RAG, LoRA/PEFT, Distillation, LangGraph, NumPy, Pandas, SQL

Web Development: TypeScript/Javascript, Next/React, Nuxt/Vue, Node, Supabase, PostgreSQL, Django, Flask

Systems: Docker, AWS EC2, Kubernetes, Git, CI/CD, Redis, WebSockets, REST APIs

Interests: Weightlifting, Reggaeton, Chess (mostly bullet), Pickleball, Dogs, Swimming, Hiking, Thai Food