

JAYDEN LEE

☎ (509) 316-6808 ✉ jeyl@berkeley.edu 🔗 [linkedin.com/in/jynlee7/](https://www.linkedin.com/in/jynlee7/) 🐙 github.com/jynlee7

EDUCATION

University of California, Berkeley

Expected Grad: May 2028

B.A. in Data Science, Domain Emphasis: Applied Mathematics

GPA: 3.7/4.0

Coursework: Optimization Models in Engineering, Data Structures and Algorithms, Data Engineering, Multivariable Calculus, Linear Algebra and Differential Equations

Organizations: Sports Analytics Group at Berkeley (Head of Business)

SKILLS

Languages: Python, R, Java, TypeScript, SQL — English + Korean

Machine Learning: PyTorch, TensorFlow, scikit-learn, NumPy, Pandas, Seaborn, Matplotlib

Tools: Git, Docker, Kubernetes, Flask, React, Supabase

EXPERIENCE

Cal Athletics (Contract)

February 2026 – May 2026

Data Analyst

Berkeley, CA

- Built a pipeline through **Snowflake** and **Colab** to quantify purchases and generate a dynamic fan scoring system.
- Developed clustering and supervised ML models to predict ticket purchase propensity from behavioral data.

Takeaway: Understanding the data fully accelerates the work done after.

Oakland Roots Sports Club (Contract)

September 2025 – December 2025

Business Analyst

Berkeley, CA

- Utilized **Python** and **Bart API** to develop a scoring system for fan engagement by integrating real-time transit data into a marketing dashboard, requiring predictive modeling and data pipeline optimization.
- Designed several visuals using **Matplotlib** and **Seaborn** effectively EDAing the data.

Takeaway: The mentor won't report to you, you ask and report to them.

Pacific Northwest National Lab

June 2024 – May 2025

Modeling & Simulation Intern

Richland, WA

- Developed a **target-decoy** statistical framework to model **false discovery rates (FDR)** for peptide identifications.
- Deployed GPU-accelerated proteomics workflows on the **Tahoma HPC cluster** via Slurm, configuring **CUDA** and batch processing to scale peptide identification validation across large mass spectrometry datasets.

Takeaway: Do not assume that results will be immediately correct, check if they're wrong from statistics.

Pacific Northwest National Lab

June 2023 – August 2023

Computational Biology Intern

Richland, WA

- Engineered **file mapping and parsing** pipeline to align 36 MGF spectral files across heterogeneous tool formats.
- Optimized the metaproteomics pipeline by processing **300,000+ mass-spec files** with GPU-accelerated Casanovo inference built on **TensorFlow**, reducing runtime and increasing throughput for large-scale analyses.

Takeaway: I like GitHub and want to do more open-source work!

PROJECTS

Personal Website ☞ | *Claude Code, Typescript, Vite, Git*

- Framed **multi-agent** development in **Claude Code**, routing each task to a purpose-built skill (design, code review, etc.) and running agents in isolated **git worktrees** to parallelize without merge conflicts.
- Authored machine-readable specs and a **Keep a Changelog/SemVer** release log, gating every AI-generated change behind **TypeScript** builds and automated accessibility checks to hold a 7K-line codebase consistent across dozens of agent sessions.

Datathon for Social Good (Clinical Intervention Simulator) ☞ | *scikit-learn, NumPy, Tableau*

- Architected a **Monte Carlo simulation** framework in Python using **NumPy vectorization** to model 300 patient deterioration scenarios, integrating a **Random Forest** model for real-time intervention optimization.
- Awarded **first place** in the 7th Annual Datathon at Berkeley alongside 50 teams and 200+ participants.

ChessBlitz ☞ | *Flask, Supabase, OpenRouter, React, SQL*

- Built backend infrastructure using **Flask** and **Supabase**, with **OpenRouter API** integration for **LLM-powered** hint generation.
- Developed the **TypeScript/React** frontend dashboard for teachers to manage individual classrooms at Berkeley Chess School.