

Angela Wu

(858)-319-9338 | angela_wu@berkeley.edu | angelawu.dev

EDUCATION

University of California, Berkeley

B.A. Data Science, Expected 2028

Berkeley, CA

Aug 2025 – Present

- **Relevant Coursework:** Calculus II, Foundations of Data Science, Structure and Interpretation of Computer Programs; Ongoing: Data Structures, Linear Algebra, Discrete Mathematics and Probability Theory

SKILLS

Languages: Python, SQL, Java, JavaScript, HTML/CSS, R

Libraries: NumPy, Pandas, SciPy, scikit-learn, Matplotlib, Seaborn, PyTorch, TensorFlow

Tools & Frameworks: Git, GitHub, Jupyter, LightGBM, statsmodels, REST APIs, Docker, AWS/GCP, PostgreSQL

TECHNICAL EXPERIENCE

Nook

Remote

Co-Founder, Product & Research

06/2026 – Present

- **Co-founded a productivity platform** designed to help students begin and sustain meaningful work through calm, socially supportive experiences
- Lead product design, web development, translating user problems into product flows and interactive prototypes
- Conduct user discovery and define product strategy across **technical and go-to-market decisions**

National Park Service (NPS) — ML & Data Science Project

UC Berkeley

ML Engineer / Data Scientist

03/2026 – Present

- Develop a **time-series forecasting pipeline** for avian botulism prediction using distinct environmental data
- **Integrate heterogeneous datasets** (lake conditions, species monitoring) w/ temporal alignment/preprocessing
- Engineer lagged and seasonal features and applying models including **SARIMAX and Gradient Boosting**
- Evaluate model performance across regions and analyzing feature importance for interpretability

Cofair AI Token Reallocation & Infrastructure Project

UC Berkeley

Data Science / Infrastructure Engineer

01/2026 – Present

- **Designed a normalized cross-provider** pricing dataset across OpenAI, Anthropic, and Google Vertex APIs, standardizing token usage, caching tiers, and long-context billing
- Built tooling for cross-provider cost analysis and **system-level optimization of AI workloads**
- Processed and structured heterogeneous API usage data to support **scalable analysis pipelines**
- Enabled downstream carbon accounting by organizing AI usage data for sustainability and ESG modeling

PROJECTS

ML Image Popularity Prediction — YOLO + LightGBM

09/2025 – 12/2025

- Built an image popularity classifier using YOLO-based object detection and LightGBM
- Engineered spatial and color features (3×3 grid counts, RGB mean/variance statistics) from images
- Trained and compared **Random Forest and LightGBM models**, analyzing feature importance, probability calibration, and class imbalance
- Developed an interactive web app to visualize predictions, object detections, and raw feature vectors for model interpretability and error analysis

Hybrid Rocketry Engines

San Diego, CA

Data & Telemetry Engineer

09/2024 – 09/2025

- Processed thrust and pressure telemetry into **performance visualizations** using Python
- Compared burn phases to identify anomalies and validate engine behavior
- Exhibited project at **OpenSauce 2025** and engaged with industry engineers

LEADERSHIP

CS Kickstart

Berkeley, CA

Design & Merch Lead

12/2025 – Present

- Lead of Merch & Design committee; direct weekly meetings, design program merchandise, and manage production and distribution for 80+ members

Speech & Debate

San Diego, CA

Secretary & National Competitor

10/2022 – 06/2025

- Elected team secretary, managing communication and logistics for a **110+** member team
- Qualified for the National Speech & Debate Tournament, placing **44th out of ~3000** public forum teams (top 1.5% of debaters) within two years of competition
- **Mentored novice debaters** to competitive success at Raven Forensics & NOVA Online Debate Camp