

# Ju Ho Kim

(352) 709-3825 | [juho\\_kim@berkeley.edu](mailto:juho_kim@berkeley.edu) | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## SKILLS

**Languages:** Python, SQL (PostgreSQL)

**ML & Data Science:** PyTorch, Hugging Face Transformers, Scikit-learn, LightGBM, SHAP, Model Evaluation/Calibration

**Data & Infrastructure:** Pandas, NumPy, Polars, FastAPI, Streamlit, Docker, Git/GitHub

## EXPERIENCE

### Data Science Undergraduate Researcher

Jan 2026 – Present

*Lawrence Berkeley National Laboratory*

Berkeley, CA

- Selected from a competitive applicant pool (top 3%) as an Undergraduate Researcher for ENIGMA, a 10-year DOE-funded national research project spanning multiple institutions
- Building a unified SQL data model and Dash-based exploration layer to consolidate siloed experimental datasets and standardize cross-institution data access for scientific teams

### Data Analyst Intern

Jun 2024 – Aug 2024

*Global ESG Foundation*

New York, NY

- Built and shipped a decision-ready Streamlit dashboard in Python (Pandas/Plotly) on global CO<sub>2</sub> emissions (1960–2023), translating raw data into executive-friendly metrics (country ranking, absolute volume, YoY growth) across 205 countries for ESG reporting
- Hardened the dataset by filtering non-country aggregates and standardizing country-level comparisons, then delivered a maintainable Git repository and presented results to leadership in a quarterly board meeting for continued use post-internship

### Medical Evacuation Coordinator

Nov 2020 – Apr 2022

*Sixth Division, Republic of Korea Army*

Gangwon, South Korea

- Deployed as an ambulance driver for landmine removal in a tightly controlled buffer zone between North and South Korea, coordinating with UN Command forces under GPS- and communication-denied conditions to deliver mission-critical medical evacuations

## PROJECTS

### Credit Risk Decisioning Platform | Python, LightGBM, Scikit-learn, SHAP, FastAPI, Streamlit, Polars, Docker

- Built an end-to-end credit decisioning system on 300K+ loan applications (2.6GB, 7-table join, 215 engineered features), applying cost-optimized approval policies that improved expected profit by 13.7%
- Served real-time default predictions via FastAPI with SHAP-based rejection explanations and actionable recourse for declined applicants, supported by a Streamlit dashboard, model drift detection experiments, and a demographic fairness audit

### News Article Political Bias Classifier | Python, PyTorch, Hugging Face, RoBERTa, NLP, Web Scraping

- Achieved 69.7% classification accuracy on 24K+ news articles spanning 13 years using fine-tuned RoBERTa transformers, with systematic ablation studies and temporal leakage checks to validate model reliability
- Engineered a reproducible NLP pipeline from raw HTML to model-ready text, designing preprocessing workflows (web scraping, normalization, deduplication of 2,147 records) to ensure end-to-end reproducibility

### Honda AURA (Honda-sponsored Multi-Agent Research Platform) | SQL, Supabase (Postgres), Next.js

- Leading the backend/data track for a Honda-sponsored multi-agent research platform, defining the Supabase/Postgres schema and row-level security policies for secure multi-tenant access control
- Implementing durable persistence for agent state and run logs, enabling reliable session recovery across multi-step autonomous research workflows

## AWARDS

### Honors in Mathematics Scholarship Award (Dean's List) | Pasadena City College

Jun 2025

Commendation for Operational Excellence | Sixth Division, Republic of Korea Army

May 2022

## EDUCATION

### University of California, Berkeley

Expected May 2027

B.A. Statistics, College of Computing, Data Science, and Society

Berkeley, CA

### Pasadena City College

Aug 2023 – May 2025

Associate of Science, Computer Information Systems

Pasadena, CA