

Leah Perez

San Francisco, CA | leah.perez@columbia.edu | leahperez.dev | linkedin.com/in/leahmperez/

EDUCATION

University of California College of the Law, San Francisco (formerly UC Hastings), San Francisco, CA

J.D. Candidate, May 2028

Relevant Coursework: Corporate Finance, Venture Capital and Start-Ups, Business Associations

Awards: CALI Legal Research and Writing 2

Columbia University, The Fu Foundation School of Engineering and Applied Science, New York, NY

M.S. in Artificial Intelligence, May 2028 (*Concentration in Finance and Operations*)

Relevant Coursework: Natural Language Processing, Artificial Intelligence

Georgia Institute of Technology, Atlanta, GA

M.S. in Computer Science, May 2028 (*Concentration in Machine Learning*)

Relevant Coursework: Machine Learning for Trading

University of California, Berkeley, Berkeley, CA

B.A. in Data Science, B.A. in Economics, B.A. in Political Science, May 2025

Relevant Coursework: Machine Learning and Data Analytics, Data Mining and Analytics, Probability for Engineers, Econometrics, Statistics for Data Science, Data Science for Economics

EXPERIENCE

Fenwick & West LLP, San Francisco, CA

Legal Engineer, May 2026 – Present; *Finance Intern (Financial Data Engineering)*, Jun. 2024 – May 2026

Oversee a portfolio of 50+ projects for the Forward Deployed Operations team, personally leading high-priority builds as PM and IC across financial, business, and legal workflows. Lead in-house data discovery for active litigation, structuring ~10TB of unstructured data and applying clustering and graph-based analysis to connect legal arguments to the evidentiary record. Develop client risk models and RAG-based drafting pipelines, and ship full-stack attorney and financial tooling, including ML-based task coding, CLM, redlining, and associate reviews. Architect the firm's core database migration to Databricks, enabling firm-wide AI and analytics.

Gusto, San Francisco, CA

Summer Associate, May 2026 – Aug. 2026

Built an AI risk assessment platform that scans source code and product documentation (GitHub repositories, PRDs, ERDs) to flag legal exposure before launch. Used a hybrid embeddings and LLM pipeline on PostgreSQL to parse statutes into machine-readable obligations, map them to risks, controls, and documented AI incidents, and match them against the codebase, routing uncertain findings to attorney review. Researched data-exposure and governance risks from agentic AI tools and reviewed commercial contracts and marketing content.

Reiner, Slaughter, and Frankel LLP, Redding, CA

Legal Intern, May 2023 – Aug. 2023

Reviewed and standardized hundreds of client narratives in mass-tort wildfire litigation, structuring damages information for case evaluation and trial readiness. Worked with attorneys and staff to redesign the claims-review workflow, improving throughput across a high-volume docket.

COMPETENCIES AND INTERESTS

Languages: Python (Pandas, NumPy, scikit-learn, PyTorch, Transformers), SQL, R, JavaScript, HTML/CSS

Tools: Flask, Databricks, PostgreSQL, Azure (Entra ID, OAuth 2.0 / OBO auth), Cloudflare, Git

Technical: ML Engineering, NLP (Language Modeling, Transformers, Fine-Tuning, RLHF), LLM Applications (RAG, Embeddings, Agentic Workflows), Reinforcement Learning (MDPs), ETL Pipelines and Data Engineering, Algorithmic Trading and Backtesting, Process Automation, API Integration, UI/UX

Statistics: Probability Theory, Regression and Econometrics, Time Series Analysis, Bayesian Inference and Bayesian Networks, Hypothesis Testing, Markov Chains, Monte Carlo Simulation

Interests: DJing and Electronic Music Production, Sewing and Designing Clothes, NYT Cooking