Evan Lee

(925)-791-9632 exl2813@berkeley.edu linkedin.com/in/evxlee evxlee.github.io

Education

University of California, Berkeley Expected Graduation: May 2029

Bachelor's in Data Science with Business and Industrial Analytics Domain Emphasis *GPA: 4.0/4.0*

Relevant Coursework: Foundations of Data Science, Structure and Interpretation of Computer Programs, Linear Algebra and Differential Equations, Principles and Techniques of Data Science, Data Structures and Algorithms, Foundations of Business

Experience

Aetheras June 2025 – August 2025

Blockchain Research Intern *Taipei, Taiwan*

- Built and deployed a personal ERC-20 token cryptocurrency on Ethereum Blockchain to understand contract behavior and token mechanics
- Conducted competitive market analysis on Blockchain's applications in the gaming industry and compared strengths, limitations, and monetization models of GameFi
- Analyzed gamer sentiment toward Blockchain and NFT integration, synthesized findings into a report on future product direction

BestBrains Feb 2025 – June 2025

Tutor *San Ramon, CA*

- Taught 50+ students (grades 1-8) in Math and English, adapting explanations to diverse learning styles
- Improved students' confidence by breaking down complex concepts and building problem-solving skills

Projects

Equity-Focused College Matcher (EquiMatch) | *Python; Tableau, Vector Space Modeling, LLM Chatbot*

- Encoded 6000+ colleges into multi-dimensional vectors (cost, diversity, location, etc.) and ranked matches using Euclidean distance with Gaussian-normalized similarity scoring
- Built a preprocessing pipeline (21,299 rows × 29 features) to handle missing data, skewed distributions, and scaled affordability/demographic indicators
- Integrated a local Llama 3.1 Chatbot for natural-language college queries; project awarded 3rd place at UC Berkeley's 7th Annual Datathon for Social Good.

Personalized Song Recommender (LyricCal) | *Python; Transformer NLP, Neural Network*

- Built an end-to-end music recommendation system that retrieves lyrics and audio features via Spotify and Genius APIs based on user-input songs
- Classified song moods using a Transformer NLP model and paired each mood label with audio features for recommendations
- Trained a Neural Network Autoencoder on acoustic features and used Cosine Similarity to identify the most similar songs for top-K recommendations

Loan Approval Predictor | *Python; Supervised Classification Models*

- Built a preprocessing pipeline with feature scaling, label encoding, outlier handling, and SMOTE class-balancing on 45k+ loan applications in the dataset
- Trained and compared multiple supervised models (Logistic Regression, Ridge Classifier, Random Forest, Neural Network) to predict loan approval outcomes
- Evaluated performance using Precision, Recall, F1 Score, and Confusion Matrices, achieving 93% overall accuracy on the final model

Technical Skills

- **Languages:** Python, SQL (PostgreSQL), R, Java
- **Libraries:** Pandas, Scikit-Learn, PyTorch, NumPy, Matplotlib, Seaborn, TensorFlow, Tableau, Hugging Face, XGBoost, LangChain, Streamlit
- **Leadership:** Samsung Solve For Tomorrow California Finalist x2 ($5000 for school), Real World Design Challenge (State 2nd Place & National Honorable Mention)