# Jonathan Ferrari

(510)-260-5474 | jonathanferrari@berkeley.edu | linkedin.com/in/jonathanferrari | github.jonathanferrari.com

## EDUCATION

**UC Berkeley** — Berkeley, CA
*Majors (BA)*: computer science, data science (honors), cognitive science, psychology — *Class of 2025*
*GPA: 4.0/4 | Units: 238.2 | Data Science Honors Degree | Highest Distinction (Summa Cum Laude)*: All Majors
Organizations: Data Science Modules, Data Science Undergraduate Studies, Data Science Society, Sigma Alpha Pi Honor Society
Psi Chi Honor Society, Helen Wills Neuroscience Institute, Relationships and Social Cognition Lab, Center for Financial Wellness

## SKILLS

**Code Languages**: Expert in Python, Ruby, Java, SQL, RegEx, TypeScript, React, R | Proficient in Julia, VBA, Lisp | Learning C++
**Libraries**: Pandas, NumPy, TensorFlow, PyTorch, Theano, Keras, CV2, SciPy, StatsModels, SkLearn, Seaborn, Plotly, MatplotLib
**Technical Skills/Coursework**: Git, Data Structures, Advanced Calculus, Linear Algebra, Discrete Math, Advanced Statistics and
Probability, Machine Learning Theory, Computer Vision, Natural Language Processing, Clustering, Classification, Random Forests,
Linear and Logistic Regression, Neural Networks, Deep Learning, Reinforcement Learning, Software Engineering, Data Engineering
**Soft Skills**: Efficient Communication, Adaptability & Flexibility, Teamwork and Collaboration, Organizational Ability, Simple and
Creative Problem Solving, Leadership, Critical-Thinking, Effective Time Management, Self-Starter, Attention to Detail, Curiosity

## EXPERIENCE

**Technical Project Manager** — Dec 2024-Present
*UC Berkeley* — *Berkeley, CA*
- Manage 10+ teams of 3-5 developers each, using Jira to track progress, allocate tasks, and ensure timely deployment
- Develop curriculum for 90+ courses across 45+ departments, impacting 13,000+ students, aligning materials with faculty goals
- Build and maintain GitHub Action workflows, automating contributions to streamline development and improve reproducibility
- Coordinate the building of Jupyter assignments for courses with 500+ students, integrating Python into non-technical subjects
- Design data-driven assignments for Environmental Studies courses, developing coral preservation projects in French Polynesia

**Full Stack Software Engineer Intern** — May-Aug 2023
*Blackhawk Network* — *Pleasanton, CA*
- Engineer a full-stack application to handle +10,000 daily financial transactions from clients and update respective databases
- Master JavaScript, React, Next.js, Java, SpringBoot, Postman, and CORS as the tech stack and fully integrate the technologies
- Implement Server-Side-Rendering solution to improve the speed of web-based applications to better the experience of customers
- Design a scalable REST API to allow for stateful communication between the user interface, backend logic, and databases
- Create modular applications with explicit documentation for improvement in scalability, deployment, and maintenance

**Machine Learning / Data Scientist Intern** — May-Aug 2022
*Liberty Mutual Insurance* — *Boston, MA*
- Work with SQL in Snowflake, SAS, and Jupyter to efficiently query massive databases with over 60B entries for relevant data
- Employ robust distributed computation, storage, and data processing to build intuitive classification models to recommend
  concrete business strategies in unique geographical regions in the U.S. based on business and publicly available data
- Create project scaffold to increase further computational efficiency and decrease operational costs for future usage/adaptation
- Employ Partial Component Analysis, along with Agglomerative Clustering, Random Forests, and Multiple Linear Regression to
  build multiple models that predict the best business strategy by geographic region to drive profitability and growth of business

**Data Engineering Lead Undergraduate Instructor** — Dec 2022-Dec 2024
*UC Berkeley* — *Berkeley, CA*
- Developed and deployed Data 101 assignments using GitHub Actions for CI/CD, Ruby, and Jekyll for website management
- Designed data engineering pipelines with Python, Jupyter notebooks, and cloud platforms (AWS, GCP) to teach ETL, data
  ingestion, transformation, transactions, query optimization, and use structured, semi-structured, and unstructured databases
- Integrated scalable data solutions (S3, BigQuery, Spark) into curriculum, allowing students to gain industry-aligned experience.
- Automated grading workflows with GitHub Actions, Otter-Grader, and Gradescope, improving efficiency for 300+ students

## RESEARCH/PROJECTS

**Real-Time American Sign Language Translation** — August 2023-Present
- Build an end-to-end webcam-based application to translate sign language in real-time to English using Deep Learning and NLP
- Implement Video Visual Transformers to predict word-level labels from video input of ~2000 classes and 100+ different sources
- Engineer BERT additive model to augment predictions based on relative occurrence frequency with previously predicted words

**Spotify Playlist Generation** — Sept 2023-Nov 2023
- Engineer a recommendation algorithm to generate a playlist based on a given song by a user, selected from the Spotify library
- Use vector embedding to generate recommendation scores between songs based on features from Spotify API and date/artist
- Build POC web application and UI based on a limited dataset from a collection of multiple volunteeers' personal Spotify data
- Implement use of Spotify API; allowing users to sample songs within the application for manual verification of performance

## INTERESTS

**Academic:** CV, Transformers/ViTs, Deep NNs, RL, Interpretability, HCI, NLP, Causal Inference, Cognition, and Decisions
**Personal:** Photography, Hiking, Camping, Poker, Social Cognition and Bias, Performant Languages, Web Development