Sara J. Miller

Berkeley, CA | sarajmiller@berkeley.edu | [LinkedIn](#) | [sarajademiller.com](#)

EDUCATION

University of California, Berkeley   *Aug 2022 - May 2026*
B.A. in Computer Science, B.A. in Data Science   *GPA: 3.81*

Relevant Coursework: Structure & Interpretation of Computer Programs, Data Structures & Algorithms, Machine Structures (Computer Architecture), Intro to Machine Learning, Intro to Database Systems, Computer Security, Discrete Math & Probability, Principles & Techniques of Data Science, Probability for Data Science, Computational Models of Cognition

WORK EXPERIENCE

SwingVision   **Berkeley, CA**
Data Analytics Intern   *Dec 2022 - Present*
- Building a React web app to streamline 3D ball-tracking data generation from 14 camera angles, improving efficiency
- Integrating TypeScript, HTML, and CSS to combine multiple linear algebra and Python scripts into a single interface
- Analyzing real-time user tendencies of 500k+ users on Amplitude to focus marketing efforts on maximizing retention
- Labeled hundreds of tennis videos with 98% accuracy, improving the app's machine learning model performance

Sports Analytics Group at Berkeley (SAGB)   **Berkeley, CA**
Data Journalist   *Jan 2023 - Present*
- Visualizing return & serving percentages on the top 50 professional women's players using Seaborn and Matplotlib
- Publishing [data-driven articles](#) and [podcasts](#) to SAGB's HTML5 UP website to provide player performance analyses

Electrical Power Research Institute   **Berkeley, CA**
Data Science Consultant   *Sep 2023 - Dec 2023*
- Cleaned a data set of 900k+ rows in JupyterHub with Python and Pandas, handling formatting issues & missing data
- Used HoloViews to develop interactive & powerful climate change visualizations for integration into EPRI's website
- Highlighted key climate patterns, including temperature and precipitation trends, from 100 cities over 50 years

PROJECTS

Dartmouth College, *Research Assistant*
- Assisting Professor Raghav Singal to create tennis handicap system for amateur players using Bayesian modeling
- Applying ML models including logistic regression, decision trees, and XGBoost to predict the next incoming shot
- Working with a group of 5 undergraduate researchers to perform feature engineering and exploratory data analysis

Spam Email Classifier (Python, JupyterHub, scikit-learn, Pandas, Matplotlib, Seaborn)
- Developed a machine learning regression model that classified spam emails in test set with 85.3% accuracy
- Applied feature engineering and sentiment analysis to minimize model's overfitting and false positive rate to 2%

Wordnet (Java, Time Series, IntelliJ)
- Developed a browser-based tool similar to Google's Ngram Viewer that handles over 14k words and hyponyms
- Converted CSV files into graphs through building a directed graph class to navigate historical synsets of similar words

LEADERSHIP EXPERIENCE

DataStory Consulting at Berkeley
External Vice President   *Apr 2024 - Nov 2024*
- Networked with 15+ industry leaders to secure 5 high-profile speakers, boosting event attendance by 40%
- Co-organized the 6th Annual Datathon for Social Good for 160+ participants, managing logistics and data sourcing

VP of Professional Development   *Nov 2023 - Apr 2024*
- Managed a team of 6 to host 5+ professional development workshops teaching club members soft & technical skills
- Mentored members through the recruiting process, providing resume reviews & mock behavioral/technical interviews

Director of Marketing   *Apr 2023 - Nov 2023*
- Boosted social media outreach by 240% by employing new marketing techniques within the club recruitment process

2024 Women in Sports Tech, Inc. Fellow   *Jun 2024 - Aug 2024*
- Selected alongside 18 fellows from a competitive applicant pool to be mentored by leaders of the sports tech industry
- Pursued a self-created fellowship path, leveraging existing internship while gaining new insights from industry experts

SKILLS & LANGUAGES

Programming Languages: Python, Java, C, HTML/CSS, JavaScript/TypeScript, RISC-V, SQL, R, LaTeX

Libraries and Frameworks: React, Jekyll, Node.js, Express, Pandas, scikit-learn, HoloViews, Git, Bitbucket, Docker, Jira