Saga Abuodeh

San Francisco, CA

linkedin.com/in/sfabuodeh
https://sfabuodeh.github.io/my-website-portfolio/#projects

Email: sfabuodeh@berkeley.edu
Phone: Available Upon request

EDUCATION

B.A. Data Science, UC Berkeley
Quantitative Economics Emphasis

EXPERIENCE

AI Trainer, Handshake Move Fellowship | June 2025 - current
- Curating and labeling diverse datasets to improve AI model accuracy and performance across multiple domains.
- Collaborating with data scientists and engineers to iteratively refine AI algorithms, with a focus on bias mitigation and enhancing model fairness and effectiveness.

Software Engineering Fellow, Headstarter AI | June 2024 – Aug. 2024
- Completed a seven-week AI Software Engineering Fellowship focused on algorithms and AI project development.
- Built final AI project and strengthened data structures, algorithms, and coding skills *(python & java)*.
- Gained hands-on experience in AI, technical interviews, and collaborative software engineering.

Data Science Research Intern, Data Discovery: Rocketship Public Schools | Aug. 2023 – Dec. 2023
- Conducted end-to-end analysis on 4TB+ of student engagement and state exam data across 15 schools to evaluate the impact of online learning programs.
- Designed and executed scalable data pipelines for ingestion, cleaning, and analysis of large education datasets
- Modeled student academic performance linked to online learning usage, identifying critical periods of slowed progress for 2,400+ students across 4 core courses.
- Delivered actionable reports to senior data scientists, directly influencing program guideline changes to improve student outcomes without increasing instructional time.

Undergraduate Teaching Assistant, UC Berkeley Stat 20 | Jan. 2023- May 2023
- Tutored and supported 220+ students in foundational statistics and data analysis using R.
- Led weekly office hours and actively answered student questions on 24/7 online support platform.
- Helped write solution sets and provided in-class academic support to reinforce lecture content.

PROJECTS

Data Engineering Pipeline | Team Project | *PostgreSQL, MongoDB, Python, Git*
- Built and analyzed Yelp database to compare relational vs. NoSQL performance for business insights.
- Identified underrated businesses, declining performance trends, and top reviewers using query optimization techniques.
- Improved PostgreSQL query speed 10x using indexes and materialized views; evaluated tradeoffs with MongoDB for semi-structured data.

Restaurant Success Prediction | Team Project | *Python, Random Forest, Logistic Regression*
- Merged multi-source datasets including Yelp reviews, zip code demographics, and economic indicators to build a predictive model for restaurant success in Pennsylvania.
- Engineered features and ran exploratory analysis to identify patterns across location, income, and cuisine types.
- Achieved 73.5% accuracy with logistic regression and 70.5% with Random Forest.

Spam Email Classifier | *Python, K-nearest-neighbor classifier, Data Visualization, Logistic Regression*
- Achieved an 82% precision using logistic regression and feature engineering. Evaluated performance through precision, false positive rate, and precision-recall curve analysis.
- Applied cross-validation and train/test split to ensure reliable results and minimize overfitting.

EXPERTISE

Python, Java, SQL, Git, R, C++, Latex, Power BI, Excel, HTML & CSS, AWS, ETL, EDA, Data Modeling, Risk Analysis, Data visualization, Data Analytics, Research, Machine Learning, A/B testing, Data Engineering, Product Development.