Bill Shao

Santa Clara · linkedin.com/in/bill-shao · billshao@berkeley.edu · bill-shao.github.io/portfolio

Education

University of California - Berkeley
B.A. Computer Science & Minor in Linguistics; GPA: 3.95 — May 2025
Relevant Coursework: Machine Learning, Optimization Models in Engineering, Efficient Algorithms, Discrete Math and Probability Theory, Databases, Computer Security, Computer Architecture, Data Structures

Experience

UC Berkeley Course Staff & Computer Science Mentor — Berkeley, CA
Teaching Assistant & Senior Mentor — 2023 - Present

- Taught and advised students in classes on Advanced Algorithms, Probability Theory, and Computer Architecture for 50+ students.
- Taskforce Member: Developed content walkthroughs & worksheet assignments. Led exam review sessions. Developed Python PDF parser to decrease parsing time of assignments by 10+ hours.
- Hosted weekly office hours aiding students and managed course logistics alongside staff and professors.

Berkeley Undergraduate Research (URAP) — Berkeley, CA
Part-Time Student Researcher — 2022 - Present

- Researched multilingualism in Large Language Models using Tensorflow for English & Mandarin idiom recognition under Dr. Babaci-Wilhite.
- Developed machine learning models using GPT3 and GPT4 APIs in Python. Increased throughput of data querying technologies by 8% in Java.
- Increased database training set size by 50% and honed data formatting skills with MongoDB.

Princeton Horizon Research — Remote
Machine Learning Research Intern — 2021 - 2022

- Investigated facial bias in CNN systems during research internship using PyTorch under Princeton Ph.D. student Angelina Wang.
- Implemented computer vision model bias correction using image datasets. Gained experience training and diagnosing models with NumPy, Pandas, and Pytorch.
- Utilized visual design skills and data management to build informative graphs of machine learning findings with Matplotlib.

Equal Opportunity Technology — Mountain View, CA
Software Engineer Intern & Instructor — 2018 - 2021

- Managed testing and integration for Snipe-it Linux database on 200+ computers. Optimized system management time by 50% during the refurbishing phase.
- Founded algorithmic curriculum and organized workshop sessions of 100+ community members on hardware and software basics.

Research & Projects

Matrix Convolution & Polynomial Multiplication Optimizer "CS61kaChow" C, Python
Implemented multithreading and optimized caching in C to increase the throughput of mathematical computations by 10x. Additionally implemented Discrete Fourier Transform in Python.

"Building Stronger Multilingual Idiomatic Repertoire Using NLP Models" 2023.
UC Berkeley AI & Linguistics research with Dr. Babaci-Wilhite. Developed Large Language Models to detect idioms in 4+ languages. Winter 2023 Conference Presentation & Publishing.

Intel Science Fair Paper: "Correction of Racial Bias in Facial Recognition Datasets" 2021.
Built Resnet based neural network to perform statistical analysis on bias of modern AI models. Practiced large-scale data formatting and presented findings to Intel Science Symposium.

Other Work & Achievements

2019 USACO Gold Competitor (Top 8% Nationwide)

Other Coursework: Stanford CS229: Machine Learning, Stanford CS231n: CNN for Visual Recognition

Skills

Languages: C++, Python, Java, SQL, React

Technologies: Pytorch, Tensorflow, Numpy, Pandas, MongoDB, Linux Systems, Git

Interests: Kubernetes, Machine Learning, Language Processing, AWS, Data Organization