Srujan Yamali

(302) 509-8614 | srujanyamali@berkeley.edu | srujanyamali.com | github.com/srujyama | linkedin.com/in/srujanyamali | US Citizen

Education

University of California, Berkeley **May 2027**
B.S. Computer Sciences GPA: **4.0/4.0**
Relevant Coursework: Machine Learning, Data Structures, Algorithms, Linear Algebra, Calculus, Differential Equations

Experience

Mercor *San Francisco, CA*
Software Engineer Intern Aug 2025 – Present
- Engineered enterprise-scale AI/ML developer tooling powering model evaluation, safety benchmarking, and data labeling workflows across top AI labs production environments.
- Automated multimodal data ingestion, fine-tuning, and continuous deployment pipelines using `Prefect`, `MLflow`, and distributed `Kubernetes` clusters—reducing enterprise experiment-to-production cycle time by 40%.

Priestley Biolabs *Berkeley, CA*
Founding Software Engineer Sept 2025 – Present
- Architected the core AI platform enabling adaptive experimental planning, reducing trial-and-error cycles in wet-lab workflows by **60%**.
- Built scalable cloud infrastructure integrating lab sensor data, reagent metadata, and ML-driven optimization tools used by early research partners.

Children's Hospital of Philadelphia - Moustafa Lab *Philadelphia, PA*
Data Science Intern November 2023 – August 2025
- Built a high-performance **time-series analysis pipeline** for genomic recombination detection using `KernelCPD` to identify shifts in protein signal distributions, scaling to **75,000+ genomes (37 TB)** using `ruptures`, `KDTree`, and multiprocessing.
- Developed a **parallelized framework** with Python multiprocessing, enabling large-scale genomic region analysis and accelerating runtime through statistical comparisons and clustering logic.

University of Delaware – Shao Lab *Newark, DE*
Software Development Intern June 2023 – August 2024
- Developed a `PyQt6/OpenCV` application to automate analysis of **730 GB** of Drosophila video, reducing manual annotation time by **90%** and saving hundreds of hours.
- Built an ROI-tracking engine using blob tracking and centroid calculations to monitor behavioral dynamics, achieving **99.7%** accuracy in mating trial analysis, streaming real-time signals to a GUI overlay for behavior classification and role tracking.

Cornell University - Hein Lab *Remote*
Machine Learning Engineer Intern March 2024 – Aug 2024
- Applied YOLO-based object detection to automate identification and tracking of fish behaviors under predation, achieving 85%+ accuracy across 500+ hours of field video footage.
- Applied deep learning techniques, including YOLO-based object detection, to automate the identification and tracking of individual and group fish behaviors due to predation from field video data.

Projects

Real-Time Behavioral Detection and Tracking | `Python, OpenCV, PyQt6, Pandas` FlyFlirt
- Developed a production-grade computer vision pipeline to automatically detect and track Drosophila behaviors across hundreds of hours of video, reducing manual annotation by **90%**.
- Implemented real-time `OpenCV/NumPy` processing for high-throughput experiments with near–zero latency and automated labeling across thousands of frames.

Genomic Changepoint Heatmap Engine | `Python, Ruptures, Scikit-learn, Matplotlib` RedCarpet
- Created a high-performance changepoint detection engine using multiprocessing and `KDTree`-based similarity search, accelerating large-scale recombination discovery by orders of magnitude.
- Automated visualization of comparative signals via `Matplotlib` heatmaps for reproducible, large-scale genomic analysis.

Skills & Interests

Languages/Frameworks: Python, JavaScript, C/C++, Rust, Java, SQL, HTML/CSS, Node.js
Libraries/Tools: React, AWS (S3, EC2, RDS), GCP, Azure, Git, Linux, Flask, Django, Docker, MySQL, PostgreSQL, SQLAlchemy, Kubernetes, REST API, Tailwind CSS, NumPy, Pandas, LangChain, PineconeDB
AI/ML: PyTorch, TensorFlow, OpenCV, Scikit-Learn, HuggingFace
Lab Skills: Gel Electrophoresis, Polymerase Chain Reaction, Mutagenesis, Cell Culturing, Bacterial Transformation