

# Srujan Yamali

(302) 509-8614 | [srujanyamali@berkeley.edu](mailto:srujanyamali@berkeley.edu) | [srujanyamali.com](http://srujanyamali.com) | [github.com/srujyama](https://github.com/srujyama) | [linkedin.com/in/srujanyamali](https://www.linkedin.com/in/srujanyamali) | US Citizen

## Education

### University of California, Berkeley

B.S. in Computer Science

**Relevant Coursework:** Machine Learning, Computer Architecture, Data Structures, Algorithms, Discrete Mathematics & Probability Theory, Signals & Systems, Circuits & Devices, Linear Algebra, Artificial Intelligence, Efficient Algorithms

## Experience

### Mercor

#### Machine Learning Engineer

San Francisco, CA

Aug 2025 – Present

- Developed and deployed production machine learning models serving **12.4M** daily inferences, reducing end-to-end latency by **37%** and increasing downstream task accuracy by **18%** through targeted feature engineering and model optimization.
- Designed a rigorous model evaluation framework spanning offline validation and live A/B testing, cutting prediction error by **29%**, improving calibration by **0.21 ECE**, and preventing **2** high-impact regressions prior to production rollout.

### Visa

#### Software Engineer Intern

Remote

May 2025 – Aug 2025

- Developed internal LLM-powered enterprise automation tools adopted by Visa's risk and product teams, reducing manual review time by **42%** and accelerating decision turnaround from days to under **2 hours**.
- Designed and deployed an AI-generated Statement of Work pipeline that automated scoping for **10,000** annual client implementations, improving drafting accuracy by **31%** and cutting onboarding timelines by **45%**.

### Cornell University

#### Machine Learning Engineer Intern

Remote

Sept 2023 – May 2024

- Implemented YOLO-based object detection and multi-object tracking pipelines for automated identification of dynamic entities in unstructured video data, achieving **85%+** accuracy across **500** hours of real-world footage.
- Developed deep learning-based computer vision systems to detect, track, and analyze individual and group-level behaviors in large-scale video datasets, enabling reliable spatiotemporal pattern extraction under noisy, uncontrolled conditions.

### Children's Hospital of Philadelphia

Philadelphia, PA

#### Data Science Intern

Sept 2024 – May 2025

- Built a high-performance **time-series analysis pipeline** for large-scale change point detection using **KernelCPD**, identifying distributional shifts in high-dimensional signals and scaling to **75,000** sequences (**37 TB**) with **ruptures** and **KDTree**.
- Designed a **parallelized analytics framework** using Python multiprocessing to accelerate large-scale sequence analysis, reducing end-to-end runtime through optimized statistical comparisons and clustering-based pruning.

### University of Delaware

Newark, DE

#### Software Development Intern

June 2023 – August 2023

- Developed a PyQt6/OpenCV desktop application to automate analysis of **730 GB** of high-resolution video data, reducing manual annotation effort by **90%** and enabling large-scale processing infeasible with manual workflows.
- Implemented a real-time ROI tracking engine using blob detection and centroid-based motion modeling, achieving **99.7%** tracking accuracy while streaming live signals to a GUI overlay for automated classification and role attribution.

## 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