

Aakash Madabhushi

The full version is below so you don't have to open a file. Available from **December 2026**, in the **San Francisco Bay Area**, and open to relocating.

aakash.vardhan15@gmail.com · [linkedin.com/in/aakash-varadhan](https://www.linkedin.com/in/aakash-varadhan) · github.com/aakashvardhan

Last updated August 2026

Education

M.S. Applied Data Science, San José State University Expected Dec 2026

AI & Machine Learning Operations Graduate Certificate, Indian Institute of Science (IISc)

B.S. Artificial Intelligence, Illinois Institute of Technology May 2023

Experience

Instructional Student Assistant, Machine Learning · SJSU Department of Applied Data Science

Aug 2026 to present

Support machine learning instruction alongside my research work.

Graduate Research Assistant · SJSU Department of Applied Data Science Mar 2026 to present

Built an end-to-end LeRobot v3 data pipeline for SO-101 pick-and-place imitation learning, synchronizing 6-DoF joint state and action trajectories with dual 640×480 RGB streams at 30 FPS, and published 50 teleoperated episodes (49,633 frames) to Hugging Face.

Developed pre-training quality gates over Parquet metadata and H.264 video, validating episode and frame alignment, video decodability and resolution, and NaN-free action and proprioceptive tensors.

Engineered a reproducible PyTorch/CUDA training pipeline for 52M-parameter Action Chunking Transformer policies with configurable augmentation, Weights & Biases tracking, 5K-step checkpointing, and resume support, completing 30K- and 60K-step runs.

Built a resumable physical-robot evaluation pipeline for a 100-trial fixed and randomized cube-placement protocol with automatic return-to-start, stale-cache guards, placement-error measurement, and success/grasp/failure logging. Completed all 50 fixed-position rollouts at 92% success and 2.49 cm mean placement error.

Building physics-based manipulation simulations in Isaac Sim and MuJoCo for SO-ARM101 and Franka Emika Panda arms, validating control algorithms before deployment.

Software Engineer · Sapaad

Jun 2023 to Oct 2024

Built production data pipelines processing 5M+ daily rows across a multi-tenant SaaS platform using Databricks and PySpark for recommendation model feature extraction.

Optimized 10+ Spark jobs through SQL window function tuning and caching, reducing dashboard query latency by 20% across 75+ tenant instances.

Implemented a Delta Lake data quality framework with automated schema validation and anomaly detection for production ML.

Computer Vision Engineer Intern · Toothlens

Sep to Dec 2022

Developed computer vision pipelines for an orthodontic alignment assessment prototype, enabling automated detection of teeth positioning from smartphone images.

Optimized a PyTorch GAN pipeline by integrating OpenCV for geometric normalization, improving model recall by 9% and reducing image preprocessing failures.

Designed evaluation protocols to benchmark model performance across edge cases including partial occlusion and low resolution.

Research Assistant · Illinois Institute of Technology

Sep 2019 to Mar 2020

Automated tweet collection through the Tweepy API and generated FastText embeddings to identify water-related disaster events.

Trained an SVM classifier to 85% accuracy, adding GridSearchCV hyperparameter tuning for a further 7% gain.

Skills

Machine learning & robotics	PyTorch · CUDA · scikit-learn · pandas · LeRobot · ACT · SmoIVLA · Imitation learning · Teleoperation · NVIDIA Isaac Sim · MuJoCo · ROS 2 · OpenCV · Weights & Biases · Hugging Face Hub
Data & pipelines	Python · SQL · PySpark · Databricks · Spark SQL · Delta Lake · Parquet · Airflow · dbt · Snowflake · Kafka
LLM applications & retrieval	LangGraph · LangChain · Llama 3.3 · Llama 3.2 Vision · Groq · Pinecone · ChromaDB · Sentence Transformers
Services & delivery	Docker · FastAPI · GitHub Actions · AWS EC2 · React · TypeScript
Also worked with	JavaScript · SwiftUI · Vite · Chrome MV3 · Gemini Nano · Cloud Firestore · MongoDB · Redis · MySQL · Streamlit · PyTorch Geometric · RoBERTa · Grad-CAM · Cross-attention fusion · Mel spectrograms · Fine-tuning ·

Projects

Teaching a Robot Arm to Pick and Place (aakashvardhan.github.io/projects/so101-robot-learning)

92% pick-and-place success, 50 rollouts at the trained cube position

Trained a real robot arm to pick up a cube and place it from 50 human demonstrations, then built a resumable evaluation harness and scored 50 physical rollouts at 92% success.

LinkedIn Simulation with an AI Recruiting Assistant (<https://github.com/aakashvardhan/linkedin-simulation>)

17 containers across 5 service groups

A LinkedIn-style hiring platform running as 17 coordinated services, plus an AI assistant that takes a recruiter from résumé to outreach and pauses for human approval before it sends anything.

Reading Emotion from Voice and Face (aakashvardhan.github.io/media/multimodal-emotion-recognition-report.pdf)

74.6% validation accuracy, four-class CREMA-D on an actor-independent split

Matched models 30 to 50 times its size on a four-class emotion benchmark with a 1.57M-parameter audio-video network trained from scratch, then showed which facial regions and frequency bands drove every prediction.

NYC Taxi Data Platform: Batch and Streaming (<https://github.com/aakashvardhan/nyc-taxi-data-warehouse>)

The same taxi data handled two ways: a scheduled pipeline that builds dashboard-ready tables joined with weather, and a streaming one that validates trips as they arrive and publishes rolling five-minute metrics.

Reader, an Attention-Management Browser Extension (<https://github.com/aakashvardhan/attention-spanner>)

0 external API calls, because the AI runs on-device

A Chrome extension and companion iOS app that help you finish what you start. The AI runs on your own machine, so nothing you read ever leaves it.

The rest of the work (aakashvardhan.github.io/#work) is on the home page, with the code and data behind each one.