

Aarya Vasantlal

860-571-1509 | aarya.vasantlal@uconn.edu | [linkedin.com/in/aarya-vasantlal](https://www.linkedin.com/in/aarya-vasantlal) | github.com/aaryavas

EDUCATION

University of Connecticut

Storrs, CT

Bachelor of Engineering in Computer Science and Engineering

Aug. 2022 – May 2026

RELEVANT COURSEWORK

Introduction to Transformers, Introduction to Machine Learning, Introduction to Artificial Intelligence, AI for Software Engineering, Introduction to Software Engineering, Linear Algebra, Differential Equations, Multivariate Calculus, Probability, Probability for Computing, Cloud Computing

TECHNICAL SKILLS

Languages: Python, C, C++, JavaScript/TypeScript, Java, Swift, R, Rust, SQL

ML & AI: PyTorch, TensorFlow, HuggingFace, Scikit-Learn, pandas, NumPy, Transformers, Matplotlib

Focus Areas: Agentic Systems, LLMs, Reinforcement Learning, RAG, NLP, Full Stack Development

RELEVANT EXPERIENCE

AI for Software Engineering Research Fellow

May 2025 – Present

University of Connecticut

Storrs, CT

◇ *Bioinformatic Research Experiment Process Knowledge Graph Extraction (Dr. Tingting Yu, Dr. Jinze Liu)*

- In process of publishing a paper on utilizing **Agentic** systems for extracting the process of experiments in research papers.
- Visualized the process of experiments across 1000s of papers in **knowledge graphs** using **React + Vite**.
- Independently orchestrated **LLMs** to normalize assay types; developed a custom **multi-agent architecture** featuring **Ontology**, **Top K**, **RAG**, and **LLM-as-a-Judge** agents.
- ◇ *Benchmarking Autonomous Bug Report Testing for Android Applications (Dr. Tingting Yu)*
- **Leading research** on novel methods for **automating bug report testing** focusing on **evals and benchmarks** for autonomous systems.
- Approaching benchmark automation with **computer-use agents** to automate extracting APK information and evaluating the success/failure of executed reports.

AI Engineer Intern

June 2025 – January 2026

Savant

Hyannis, MA

- **Independently spearheaded** the end-to-end development of an on-device **Voice AI model**, replicating advanced conversational AI performance (e.g., OpenAI GPT-4o) with **limited computational resources**.
- Designed and implemented a **multi-agent architecture** to manage over **60 complex tool calls**.
- Developed a local pipeline integrating **OpenAI Whisper** for Speech-to-Text and a custom **model2vec embeddings + classifier** to intelligently route queries to **lightweight SLM Agents**.
- Engineered a sophisticated **semantic router** to categorize tool-call queries, directing them to specialized agents.
- Built a **vector database for chat history** leveraging **RAG** for relevant context fetching.
- **Finetuning** and **training** weights on custom **LLMs, CSMs** for up to **95% accuracy per use case** and **human-like speech**.

Full Stack Engineer Intern

January 2025 – June 2025

Flashpoint AI

New York, NY

- Built a full stack map query application as a solo project for spatial survey data visualization.

- Database with **PostgreSQL** and **PostGIS**; backend and API with **Node.JS**, **Next.JS**, and **TypeScript**; frontend with **React** and **TailwindCSS**.
- Implemented an **OpenAI API** agent for translating natural language user input to **PostGIS** geo-spatial queries.

Machine Learning Research Fellow

September 2024 – August 2025

EL GATO Lab

Storrs, CT

- ◊ ***Predicting Motion Outcomes for Civil Court Cases (Dr. Derek Aguiar)***
- Applied feature engineering to tabular data (**Beta Bernoulli**, **One Hot Encoding**) and large document data (**TF-IDF vectorization**) for better prediction accuracy.
- Combined **TF-IDF** weights with **Word2Vec** embeddings to create weighted embeddings for classification.
- Applied classifiers from **scikit-learn** including **Decision Tree**, **XGBoost**, and **SVM** for up to **65%** prediction accuracy on motions.
- Explored **anomaly detection** methods such as **isolation forests** and **One Class SVM** for data filtering.

AI RLHF Trainer for Mathematics

May 2024 – August 2024

Outlier

San Francisco, CA

- Engineered high-quality alignment data for **Frontier Models** for complex tasks such as **Differential Equations** and **Linear Algebra**.
- Identified and mitigated **surface-level agreeability** by engineering adversarial feedback loops prioritizing **rule-based structures** and mathematical correctness.
- Evaluated model outputs for **deceptive reasoning**, contributing to the safety and human-alignment of data.

PROJECTS

Multi-Objective RL Food Recommendation System | *Python, RL*

September 2025 – April 2026

- ◊ ***Improving Healthiness in MOPI-HFRS (Research Paper) using A2C RL (Dr. Chuxu Zhang)***
- **Won First Place in UConn's School of Computing College of Engineering Senior Design Competition (1 of 72)**
- Extending a static **Graph Neural Network** food recommender into a dynamic, **multi-objective** system using **Advantage Actor-Critic RL** to jointly optimize **relevance** and **health alignment** across 8,170 users and 6,769 food items.
- Discovered and corrected **3 compounding evaluation bugs** in the original codebase that misreported model accuracy by over 50%, establishing a corrected performance baseline.
- Replaced a high-variance REINFORCE policy gradient with an **A2C critic baseline** and added **behavioral cloning pretraining** to warm-start the policy before RL fine-tuning.
- Designed a **per-step multi-objective reward** combining relevance and health tag overlap, boosting health alignment from **0.39 to 0.73** while leading to slight decrease in preference alignment.

Terminal Based LLM Trainer | *LLMs, HuggingFace, Python*

September 2025 – December 2025

- Developed a full **TUI** tool to automate **data generation**, **finetuning**, and **alignment** training for non-ML background engineers.
- Employed **VB-LoRA** for efficient finetuning of models of various sizes on consumer hardware.
- Iteratively aligns model outputs using **ORPO**, contrasting ground truth against the model's own prior generations to progressively steer behavior without a separate **RLHF** stage.

Neural Network from Scratch | *Deep Learning, Python, NumPy*

May 2024 – August 2024

- Implemented a **feed forward neural network** without deep learning frameworks for handwritten digit detection.
- Developed core components including **forward propagation**, **backpropagation**, and **gradient descent**.