

# Gautham Kishore

650-889-9370 | [gkishore@ucsd.edu](mailto:gkishore@ucsd.edu) | [linkedin.com/in/gauthk6](https://www.linkedin.com/in/gauthk6) | [github.com/gauthk6](https://github.com/gauthk6)

## Education

### University of California, San Diego

La Jolla, CA

BS, Mathematics-Computer Science **Minor:** Data Science

Sep. 2022 – June 2026 (GPA: 3.8)

- **Quantitative Coursework:** Machine Learning Algorithms (Graduate), Deep Reinforcement Learning, Deep Learning, Statistical Learning, Intro to Probability, Design & Analysis of Algorithms, Advanced Data Structures.

## Research & Professional Experience

### Research Assistant

June 2024 – Present

MixLab, University of California San Diego

La Jolla, CA

- Leading development of self-aware autonomous research agent that maintains explicit representation of its knowledge state, hypotheses, and research gaps, performing competitively with leading proprietary systems (Gemini Deep Research, OpenAI Deep Research) while achieving superior cost efficiency; on track to be among top open-source implementations.
- Architected metacognitive system enabling dynamic query refinement and strategic planning through structured knowledge graphs, vector databases for semantic retrieval, and iterative hypothesis evaluation with novel stopping criterion; preparing first-author publication.
- Co-authored ACL Findings 2025 paper on Vision-Language Model limitations; invented Standardized Relative Entanglement (s-RE) metric to quantify model sensitivity to input perturbations, enabling diagnosis of representation errors and spurious correlations through 150+ controlled experiments.

### AI Engineer Intern (Part-Time)

Sep. 2025 – Present

Stealth AI Startup

Remote

- Building end-to-end search agent replicating reasoning patterns of frontier models (Gemini, GPT-4) for Generative Engine Optimization (GEO) applications; system decomposes queries into parallel search paths, retrieves and reranks relevant content, and generates responses matching target model citation behavior.
- Engineered multi-stage retrieval pipeline with embedding-based first-pass filtering and fine-grained reranking, integrated with open-source LLMs and asynchronous architecture to achieve production-grade latency while maintaining high source attribution accuracy.

### Data Science and ML Intern

June 2024 – March 2025

Legion Technologies

Santa Clara, CA

- Architected scalable data extraction pipeline using GPT-4 API to scrape and structure academic calendar data from university websites, creating new predictive features for demand forecasting models.
- Built robust data validation and cleaning pipeline for ESPN API sports schedule data across 10,000+ events, improving reliability of critical input signals for workforce management models.

### Machine Learning Research Intern

July 2021 – Sep. 2022

University of Waterloo

Waterloo, ON

- Co-authored ECAI-2024 paper on neural code translation; developed reinforcement learning-based tool for Java/Python/C++ program translation using compiler and symbolic execution feedback, achieving state-of-the-art accuracy across 57,000 code pairs.

## Projects

### Multimodal Retrieval Augmented Generation (RAG) for Education | Python, PyTorch, GenAI

Jan. 2025

- Architected dual-path RAG framework combining text and image retrieval for multimodal question answering in biology and chemistry, validating that processing both modalities is crucial for complex scientific reasoning.
- Achieved 98% accuracy with GPT-4o on 50-question college-level evaluation dataset, representing 42 percentage point improvement over text-only baseline and nearly eliminating insufficient information responses.

## Technical Skills

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

**Libraries & Frameworks:** PyTorch, TensorFlow, LangChain, Scikit-learn, Pandas, NumPy, Hugging Face, OpenAI API

**Tools & Technologies:** Git, GitHub, Docker, Vector Databases, Search APIs