

# Karthik Thyagarajan

karthik6002@gmail.com | kthyagar@purdue.edu

 github.com/karthikcsq |  linkedin.com/in/karthikthyagarajan06 | www.karthikthyagarajan.com

## EDUCATION

### Purdue University

B.S. of Computer Science & Artificial Intelligence - 4.0 GPA

August 2024 - May 2027

West Lafayette, Indiana

**Relevant Coursework:** Data Structures and Algorithms, Computer Architecture, Programming in C, Linear Algebra

## SKILLS

- **Languages & Frameworks:** Python, Java, C++, C, JavaScript/TypeScript, React, Next.js, Flask, FastAPI
- **Data Engineering & Databases:** PostgreSQL, NoSQL (Firestore), Vector Databases (Pinecone), ETL Pipelines, Data Modeling, Query Optimization
- **AI/ML & Data Science:** PyTorch, TensorFlow, LangChain, RAG Systems, NumPy, Pandas, Computer Vision, NLP
- **Cloud & DevOps:** AWS (S3, Lambda, EC2), GCP, Docker, CI/CD, Git, Linux, REST API

## EXPERIENCE

- **Machine Learning Engineering Intern** Jun 2025 - Present  
Peraton Labs (Internship & Part-Time Co-op) Silver Spring, MD
  - Engineered a reinforcement learning agent for IoT malware detection, achieving 35% reduction in exploration latency and 25% increase in detection coverage over brute-force approaches through custom reward shaping and policy optimization.
  - Designed heterogeneous graph neural network architecture with autoencoders to model inter-device communication patterns, accelerating RL convergence by 40% and improving anomaly detection accuracy by 18%.
  - Implemented distributed training pipeline using PyTorch and AWS EC2, processing 500K+ IoT device events daily.
- **Computer Vision Software Engineer** Feb 2025 - Aug 2025  
Memories.ai (Part-Time) Remote
  - Architected and deployed production video memory framework processing 10K+ video streams for AR applications, implementing efficient data pipelines with 60% throughput improvement through frame sampling and caching optimizations.
  - Developed and published Python SDK (pymavi) on PyPI with 2K+ downloads, featuring comprehensive API wrappers, async processing, and developer documentation for video analysis workflows.
  - Built RESTful microservices with Flask and PostgreSQL for spatial data indexing, reducing query latency by 45%.
- **Undergraduate Robotics Researcher** Mar 2025 - Jun 2025  
IDEAS Lab, Purdue University (Part-Time) West Lafayette, IN
  - Developed real-time SLAM pipeline in C++ and Python integrating sensor fusion algorithms, improving 3D reconstruction accuracy by 25% and reducing mapping errors through Kalman filtering and loop closure detection.
  - Optimized navigation algorithms with multithreading and memory management techniques, decreasing computational overhead by 30% and enabling real-time processing on embedded systems.
  - Implemented novel view synthesis using neural radiance fields (NeRF), generating photorealistic scene reconstructions for robotic path planning.
- **Undergraduate Data Engineer** Aug 2024 - Dec 2024  
The Data Mine Corporate Partners, Purdue University (Part-Time) West Lafayette, IN
  - Built end-to-end weed detection data pipeline processing 200GB+ drone imagery, implementing ETL workflows with Python, TensorFlow, and PostgreSQL with indexed queries for 40% faster retrieval.
  - Engineered semantic segmentation models (U-Net, DeepLabv3) achieving 92% accuracy on 50K+ labeled images, reducing herbicide usage by 60% and operational costs by \$150K annually.
  - Designed normalized database schema with spatial indexing for geolocation data, optimizing JOIN operations and supporting real-time analytics dashboard for field teams.
- **ML Science & Engineering Apprenticeship** Jun 2023 - Aug 2023  
Naval Research Laboratory (Full-Time) Washington, D.C.
  - Led 4-engineer team building deep learning models (UNet, Transformers, GANs) for underwater acoustic modeling, improving transmission loss prediction RMSE by 20% over physics-based simulations through ensemble techniques.
  - Architected secure RAG system with vector embeddings and access controls for classified document retrieval, implementing LangChain with custom document loaders and reducing query response time by 65%.
  - Developed data preprocessing pipelines handling 5TB+ acoustic datasets, implementing parallel processing with multiprocessing and Dask for 10x throughput improvement.

## PROJECTS

---

### • Frontera (Founding Engineer)

Ongoing  
[frontera.app](https://frontera.app)

Tools: Next.js, TypeScript, FastAPI, Python, LangChain, Firebase, NoSQL, REST API

- Building full-stack platform as founding engineer for AI-powered project management, architecting microservices backend with FastAPI, Firebase Authentication, and NoSQL database serving 500+ beta users.
- Engineered multi-agent LLM system using LangChain with custom tools for roadmap generation, task decomposition, and intelligent issue resolution, reducing project planning time by 70%.
- Implemented real-time WebSocket connections and event-driven architecture for collaborative workspace features, handling 1K+ concurrent sessions with 99.9% uptime.

### • Caladrius (AI Medical Triage)

Sep 2025

Tools: React Native, Python, LangGraph, GPT-4, AWS S3, Cryptography, REST API

[github.com/karthikcsq/Caladrius](https://github.com/karthikcsq/Caladrius)

- Architected cross-platform mobile application with React Native integrating encrypted QR-based medical data transfer, implementing AES-256 encryption and zero-knowledge architecture for HIPAA-compliant data handling.
- Built multi-agent diagnostic system with LangGraph processing patient symptoms through specialized agents (history analysis, symptom extraction, differential diagnosis), achieving 85% triage accuracy; awarded 2nd Place at HackGT 12.
- Designed RESTful API backend with Python and AWS S3 for secure medical record storage, implementing rate limiting and OAuth 2.0 authentication.

### • Personal Portfolio & Blog Platform

Ongoing

Tools: Next.js, TypeScript, React, Tailwind CSS, Pinecone, AWS S3, Python, Vercel

[github.com/karthikcsq/personalsite](https://github.com/karthikcsq/personalsite)

- Developed full-stack portfolio site with Next.js and TypeScript deployed on Vercel with AWS S3 CDN integration, achieving 95+ Lighthouse performance score and <1s load times globally.
- Implemented RAG-powered semantic search using Pinecone vector database for technical blog content, processing 100+ markdown documents with GPT-4 embeddings for intelligent content discovery.
- Built automated CI/CD pipeline with GitHub Actions for continuous deployment, image optimization, and content validation.