

KARTHICK T. SHARMA

Brisbane, Australia

karthicksharma11@gmail.com || [linkedin.com/karthick47](https://www.linkedin.com/in/karthick47) || github.com/Karthick47v2

Education

MPhil – Queensland University of Technology, Brisbane, Australia Exp. Feb 2028

ANZCORS Group Scholarship Recipient | AI for Pediatric Cardiac Surgery Outcomes

B.Sc. Computer Engineering – University of Sri Jayewardenepura, Colombo, Sri Lanka July 2024

Minor in High-Performance Computing | CGPA: 3.86/4.00

Experience

Co-Founder & CTO – Stealth Health-Tech Startup, USA (Remote) Oct 2025 – Present

- Leading a team of engineers to architect and develop a medical agentic AI system leveraging foundation models and multimodal AI for real-time clinical workflows

Machine Learning Research Engineer – Neuralgap, USA (Remote) July 2024 – Dec 2025

- Developed a multimodal foundation model for structure-based drug design, securing **early-stage industry funding**
- Reduced clinician wait time for cardiac video analysis from **7 minutes (legacy system) to 3 seconds** by deploying a TensorRT-optimized 3D CNN-Transformer, enabling faster patient diagnosis with **1.8% improved accuracy**
- Designed multimodal PyTorch models and developed Flask APIs for a clinical decision support platform
- Built an agentic AI interpretability platform for drug discovery, leading a junior engineering team

Research Assistant – Nanyang Technological University, Singapore (Remote) Nov 2023 – Aug 2024

- Deployed a semi-supervised anomaly detection system at a major steel manufacturer, boosting rare equipment fault detection (1 in 1,000) to 68.3% (**+33% over legacy**) and **replacing manual inspections** via continual adaptation
- Conducted preliminary experiments on improving answer relevance in RAG, helping secure **SGD 180,000 grant funding**

Research Intern – Singapore University of Technology and Design, Singapore Feb 2023 – Aug 2023

- Built a self-updating sentiment analysis system for fashion retail using Apache Flink and Kafka, maintaining **85.5% accuracy over 8 years** of historical trend shifts, processing 3,511 reviews/sec with **zero manual maintenance**
- Developed distributed benchmarking pipelines for LLMs (T5, GPT-2, FLAN-T5) with PyTorch and DeepSpeed to evaluate factual knowledge retention in traditional continual learning techniques, tracked with Weights & Biases

Skills

Languages: Python, C++, C#, JavaScript

ML/DL: PyTorch, TensorFlow, Scikit-learn, LLMs, RAG, Computer Vision, NLP

MLOps & DevOps: Docker, GitHub Actions, Weights & Biases, GCP, Azure

Development: Flask, FastAPI, React, Streamlit, Flutter

Databases: MongoDB, MySQL

Projects

Automated MCQ Quiz Generator | *Python, Dart, PyTorch, FastAPI, Flutter, Docker, GCP, Firebase* [Source](#)

- Fine-tuned T5 for MCQ generation, accelerated inference 2.3× via ONNX and deployed on GCP with FastAPI
- Improved question quality by generating diverse, relevant distractors using spaCy, KeyBERT, and Word2Vec
- Implemented Dockerized backend with CI/CD pipelines on GitHub Actions for automated testing and deployment

Virtual Personal Fitness Assistant | *Kotlin, Python, C#, TensorFlow, Android Studio, Unity, Firebase* [Source](#)

- Developed a real-time exercise posture correction Android app utilizing a TF-Lite optimized pose estimation model
- Finalist** at Mozilla Campus Club of IIT's Open Hack 2021 for innovative use of real-time ML on edge devices

Publications

[1] Sharma, K., et al. *ED-Triage-Agent: A Framework for Human-AI Collaborative Emergency Triage*. medRxiv 2026.

[2] Sharma, K., et al. *Enhancing Medical Diagnostics via Unpaired Multimodal Data Integration*. ISBI 2025.

[3] Zhao, D., Sharma, K., et al. *Scalable ML for Real-Time Fault Diagnosis in Industrial IoT Systems*. ICDE 2025.

[4] Wu, Y., Sharma, K., et al. *SentiStream: Co-Training for Adaptive Online Sentiment Analysis*. EMNLP 2023.

[5] Shi, T., Wu, Y., Sharma, K., et al. *Online Continual Knowledge Learning for LLMs*. arXiv 2023.