

Kamalasankari Subramaniakuppusamy

iamkamalamskls@gmail.com | +1 (571) 307-8802 | bit.ly/kamalagithub | bit.ly/kamalalinkedin | bit.ly/kamalaportfolio | Google Scholar

RESEARCH SUMMARY AND INTERESTS

My research explores the reasoning capabilities and emergent behaviors of large language models (LLMs), with a focus on understanding and mitigating their failure modes. I investigate fundamental questions of LLM cognition—why models hallucinate, how structured prompting and chain-of-thought mechanisms enhance reasoning, and what enables emergent behaviors to arise from pre-trained systems. By bridging theory and application across code understanding, natural language processing, graph theory, and Agentic AI, my work advances the foundations of reliable, interpretable, and self-correcting AI agents capable of transparent and adaptive reasoning.

Keywords: Large Language Models (LLMs); Reasoning and Cognition; Emergent Behaviors; Chain-of-Thought; Agentic AI; Explainable AI; Graph-Based Learning; Code Understanding; AI Reliability

EDUCATION

Master of Science in Computer Science

The George Washington University, *Washington, DC*
GPA: 3.5 / 4.0

May 2026 (Expected)

Bachelor of Engineering in Electronics and Communication Engineering

St. Joseph's College of Engineering, *Chennai, India*
CGPA: 9.3 / 10

May 2024

PUBLICATIONS

Bridging Semantics & Structure for Software Vulnerability Detection using Hybrid Network Models

14th International Conference on Complex Networks and Their Applications

Dec 2025

- Introduced a hybrid GNN + Transformer framework combining structural and semantic cues for vulnerability detection.
- Achieved superior precision on multi-language datasets, demonstrating synergy between graph topology and LLM embeddings.
- Provided interpretable graph-attention visualizations linking vulnerable code regions to learned semantic patterns.

SecureFixAgent: A Hybrid LLM Agent for Automated Python Static Vulnerability Repair

2025 International Conference on Machine Learning and Applications (ICMLA)

Dec 2025

- Designed a pipeline integrating LLM reasoning with multi-armed bandit optimization for autonomous code repair.
- Enhanced patch success by 13.5% and reduced false positives by 10.8% over existing repair baselines.
- Demonstrated explainable patch rationales through language-grounded reasoning traces.

MLCPD: A Unified Multi-Language Code Parsing Dataset with Universal AST Schema

arXiv Preprint

Oct 2025

- Released the first cross-language dataset unifying AST representations across 10 programming languages.
- Developed a standardized JSON–Parquet conversion pipeline for scalable code structure analysis.
- Enabled benchmarking of language-agnostic reasoning, summarization, and repair models.

MalCodeAI: Autonomous Vulnerability Detection and Remediation via Language-Agnostic Code Reasoning

IEEE 26th International Conference on Information Reuse and Integration (IRI 2025) Aug 2025

- Proposed a two-stage LLM architecture for code comprehension, exploit simulation, and automated patching.
- Incorporated CVE-based risk scoring and self-verification loops to improve detection reliability.
- Demonstrated a generalizable framework for autonomous vulnerability reasoning across languages.

YOLO: Roof Material Detection Using Aerial Imagery

2024 International Conference on Computing and Data Science (ICCDSS) Apr 2024

- Developed a YOLOv7-based aerial-image classifier improving mAP by 32.52% over YOLOv5.
- Implemented high-resolution preprocessing using CARTOSAT-3 imagery for urban infrastructure mapping.
- Awarded *Best Paper* for advancing smart city resilience through remote-sensing analytics.

RESEARCH PROJECTS

VulnGraph: Graph + LLM Embeddings for Vulnerability Detection Jul 2025 – Sep 2025

- Developed a multimodal fusion model integrating AST/CFG graph structures with LLM-based semantic embeddings through a two-way gating architecture.
- Achieved 93.57% accuracy—surpassing GNN-only (+8.36%) and LLM-only (+17.81%) baselines—while producing explainable saliency subgraphs and natural language justifications.
- Published at *Complex Networks 2025* (Springer SCI Series, SCOPUS-indexed).

SecureFixAgent: Hybrid LLM Agents for Vulnerability Detection May 2025 – Aug 2025

- Designed a hybrid agentic framework combining LLM reasoning with Bandit optimization for continuous vulnerability detection, repair, and validation.
- Fine-tuned open-source LLMs using LoRA on Apple MLX and NVIDIA CUDA, improving patch success by 13.5% and reducing false positives by 10.8%.
- Presented at ICMLA 2025 under the “Robustness and Security of LLMs” special session.

MalCodeAI: Autonomous Malicious Code Detection Jan 2025 – May 2025

- Built a two-phase LLM pipeline using a fine-tuned Qwen2.5-Coder-3B-Instruct model for semantic code reasoning and vulnerability repair.
- Integrated exploit simulation, CVE-based risk scoring, and automatic patch generation for explainable code remediation.

MLCPD: Multi-Language Code Parsing Dataset May 2025 – Oct 2025

- Created MLCPD, a cross-language dataset covering 10+ programming languages with a unified AST schema for structured code representation.
- Implemented a JSON–Parquet data pipeline enabling scalable AST normalization and graph-based program learning.

- Facilitated language-agnostic reasoning tasks such as vulnerability detection, summarization, and repair.

YOLO: Roof Material Detection Using Aerial Imagery

Jul 2023 – Apr 2024

- Led a deep learning system classifying metallic vs. concrete roofs from aerial imagery, improving mAP by 32.52% over YOLOv5.
- Employed CARTOSAT-3 data and GPU acceleration for real-time urban analytics and disaster preparedness.
- Awarded *Best Paper* at IEEE ICCDS 2024 for impact in smart city infrastructure.

RELEVANT EXPERIENCE

Student Academic Assistant

The George Washington University, *Washington, DC*

Sep 2025 – Present

- Assisting instruction for Advanced Machine Learning (CSCI 4907 / 6365) through grading, course refinement, and office hours.
- Supporting faculty and students on deep learning, probabilistic modeling, optimization, and graph-based learning topics.

Software Engineer Intern

Lillup, *Remote, USA*

Apr 2025 – Aug 2025

- Developed AI-driven Voice Tutoring and Research Assistant systems for the Talent Passport platform across all iOS versions.
- Engineered modular data pipelines for JSON, XML, CSV, and Parquet to automate cross-format data ingestion and synchronization.

CERTIFICATIONS & HONORS

GW SEAS Dean's Award Scholarship

2024

Best Paper Award, IEEE ICCDS

2024

TECHNICAL SKILLS

Programming: Python, C++, Java, Bash, SQL, R, MATLAB, JavaScript

Machine Learning: LLMs, Deep Learning, Transformers, NLP, Code Understanding, Computer Vision, Graph Neural Networks (GNNs), Anomaly Detection, Multimodal Learning, Explainable AI, Agentic AI

Frameworks: PyTorch, TensorFlow, Scikit-learn, Hugging Face, LangChain, Ollama, PyG, OpenCV, YOLOv5, Pandas, NumPy

Tools & Platforms: Git, Docker, Spark, Hadoop, Apple MLX, AWS, Google Cloud, Linux/Unix

Professional: Research Writing, Technical Communication, Mentorship, Academic Presentation, Project Leadership