

Kushal Raj G S

☎ 9686293233 | ✉ kushalrajgs@gmail.com | 📍 Bengaluru, India | 🌐 Kushal-Raj-G-S | 📱 Kushal Raj G S | 🌐 kushalrajgs.in

Education

B.E. Artificial Intelligence and Machine Learning, CGPA: 8.98

2023 – 2027

BMS Institute of Technology and Management, Bengaluru

Technical Skills

AI & LLM Engineering: Agentic AI, Multi-Agent Systems, CrewAI, LangChain, Retrieval-Augmented Generation (RAG), Model Context Protocol (MCP), Knowledge Graphs, Semantic Routing

Computer Vision & Machine Learning: Computer Vision, MediaPipe, DSP, CNNs, SentenceTransformers, FAISS, GLiNER

Full-Stack Engineering: Python, FastAPI, Node.js, Next.js, React, TypeScript, PostgreSQL, REST APIs, WebSockets

Infrastructure & Developer Tools: Docker, Supabase, Neon PostgreSQL, Cloudflare R2, Git, GitHub, Claude Code / Cowork

Projects

Roast – AI Review Intelligence Platform

roast.systems | 🌐 | 2025 – 2026

- Scaled a full-stack review intelligence platform to process **200K reviews**, filtering **149K+ noisy reviews**, generating **50K+ 384-dimensional embeddings**, and transforming large-scale customer feedback into actionable engineering insights with **20 prioritized issue clusters**, while reducing **timeout failures by 40% across large-scale bulk processing workloads**.
- Devised a **semantic issue clustering pipeline** using **vector embeddings** and **FAISS** similarity search to identify recurring product issues and generate AI-assisted **Root Cause Analysis (RCA) reports**, reducing **embedding API calls by 98%**.
- Orchestrated a **hybrid multi-version shadow deployment framework (v1/v2/v3)** executing comparative inference across model versions, enabling **safe rollouts, drift monitoring, performance regression analysis**, and **100% output consistency validation** through **asynchronous processing pipelines** and a **fault-tolerant production architecture**.

Baxel - AI Powered Spec-to-Backend System

baxel.live | 🌐 | 2026

- Assembled a production-grade **AI backend generation platform** transforming natural-language specifications into **6 clean-architecture backend artifacts** through a **parallel dual-flow NLP pipeline**, **RAG-powered blueprint retrieval**, **Pydantic-enforced structured generation**, and **hierarchical multi-agent reasoning**.
- Synthesized a **hierarchical AI reasoning pipeline** combining **GLiNER zero-shot entity extraction**, **Semantic Router**, **SentenceTransformers**, **Faiss semantic clustering**, and an **Architectural RAG engine** over a **150-blueprint vector database** to normalize specifications, resolve entity ambiguity, and improve generation consistency.
- Streamlined a **hierarchical multi-agent architecture** where a **70B Principal Architect** generated production-ready system designs while **DBA, Security, PM, and Anti-Fragility** agents executed concurrent design reviews in **5–10s**, alongside **SAGE**, a specification-aware chatbot with **5-turn conversational memory** and grounded architectural reasoning.

Varinth – Agentic AI Verification Platform

🌐 | 2026

- Architected a **proof-first AI verification platform** integrating the **Model Context Protocol (MCP)**, **hybrid lexical-semantic retrieval**, **retrieval-augmented verification (RAV)**, and **1024-dimensional semantic embeddings** to transform LLM responses into **auditable Proof Objects** with grounded evidence, file paths, and line-level citations.
- Engineered a **3-agent Critic–Verifier–Judge verification swarm** performing **claim decomposition**, **evidence grounding**, **contradiction detection**, **Grounded Correction** generation, and deterministic verdict synthesis through **parallel asynchronous execution**, **adversarial multi-agent reasoning**, and **evidence-driven consensus** for **trustworthy AI verification**.
- Built a secure **memory-augmented verification architecture** supporting **autonomous answer generation** and **self-verification**, **semantic memory caching**, **knowledge graph reasoning**, **NVIDIA NemoClaw/OpenClaw sandbox policies**, and bounded **500-file verification scopes** to enable safe repository-wide verification with optimized repeated audits.

Experience

AI Intern - VITAL | FinArna Technologies Pvt. Ltd.

🌐 | Feb 2026 – Aug 2026

- Architected **VITAL (Vision-based Intelligent Triage and Autonomous LifeSign analytics)**, an **AI-powered contactless clinical triage platform** integrating **computer vision**, **biomedical signal processing**, **multi-agent AI** and **LLM-driven clinical reasoning** to estimate **6 physiological metrics** from RGB video for **real-time** emergency decision support.
- Engineered an end-to-end physiological inference pipeline combining **468-landmark MediaPipe face tracking**, adaptive **250 Hz rPPG reconstruction**, **POS projection**, **Butterworth filtering**, **Welch PSD**, and **FFT** for robust contactless vital-sign estimation supporting **multi-agent clinical AI orchestration** under real-world lighting, motion, and head-pose variations.
- Designed an **agentic clinical decision-making system** using a **3-agent CrewAI workflow (Perception, Diagnostic, and Coordinator)** integrated with **NVIDIA NIM** to perform **multi-agent reasoning**, **ESI Level 1–5 patient prioritization**, **compensated shock detection**, and drive **ARIA multimodal assistant** with OCR-enabled clinical document understanding.

Achievements and Recognition

- Secured **5 national hackathon victories**, including **1st Place**, **1st Runner-up**, and multiple **Top-3 finishes**, by building AI-driven automation and intelligent decision-support systems.
- Ranked among the **top 1%** in Smart India Hackathon, leading development of scalable AI-driven solutions later published and presented at an **IEEE conference**, earning the **Best Paper Award** for real-world impact and innovation.