

Roger Richard Demello

Nagpur, Maharashtra, India | +91 94217 78898 | rogerdemello289@gmail.com
linkedin.com/in/rogerdemello | github.com/rogerdemello | rogerdemello.tech

SUMMARY

AI/ML and Software Engineer building and deploying **production-grade AI systems** - large language model (LLM) applications, Retrieval-Augmented Generation (RAG) pipelines, autonomous multi-agent systems, and end-to-end machine learning workflows. Shipped **3+ production applications** serving **200+ users**, cutting response latency by **35%**, lifting task-completion efficiency by **40%**, and accelerating delivery cycles by **50%**. Fluent across the full stack and the full ML lifecycle - feature engineering, model training, evaluation, deployment, observability, and monitoring - on a strong computer-science foundation of data structures, algorithms, and system design. B.Tech graduate from RCOEM with a Minor in Artificial Intelligence & Machine Learning.

TECHNICAL SKILLS

Programming Languages: Python, C++, Java, C, TypeScript, JavaScript, SQL, Git

Generative AI & LLMs: Retrieval-Augmented Generation (RAG), Prompt Engineering, Agentic AI, Multi-Agent Systems, Embedding Models, Semantic Search, LangChain, OpenAI, Azure OpenAI, Hugging Face Transformers

Machine Learning & Deep Learning: Scikit-learn, PyTorch, Random Forest, XGBoost, IsolationForest, Anomaly Detection, NLP, Computer Vision, Feature Engineering, Model Evaluation, Cross-Validation, statsmodels

Data & MLOps: Pandas, NumPy, SciPy, joblib, ONNX / onnxruntime, Model Deployment, MLOps, Git LFS Model Versioning

Vector Databases & AI Infrastructure: FAISS, Pinecone, ChromaDB, AI System Design

Backend & APIs: FastAPI, Flask, Express.js, Node.js, REST APIs, WebSockets, Prisma ORM, SQLAlchemy, Pydantic

Frontend: React, Next.js, TypeScript, Tailwind CSS, HTML, CSS

Databases: PostgreSQL, Supabase, Redis, SQLite, MySQL

DevOps & Cloud: Docker, CI/CD, GitHub Actions, AWS, Vercel, Render, Gunicorn, Sentry, Kubernetes-ready

CS Fundamentals: Data Structures, Algorithms, Object-Oriented Programming

PROFESSIONAL EXPERIENCE

AI LifeBOT

Jan 2026 - July 2026

AI Engineer Intern

Remote

- Engineered **3 production LLM applications** and autonomous AI agents (LangChain, GPT-4) serving **200+ users**, increasing task-completion efficiency by **40%** through tool integration, long-term memory, and goal-driven agentic workflows.
- Architected end-to-end **Generative AI pipelines** (RAG, vector databases, external APIs, real-time streaming), reducing response latency by **35%** and sustaining **99.5% uptime**.
- Designed **multi-agent orchestration** with approval gates, telemetry pipelines, and audit logging for reproducible agent evaluation.
- Led prototyping of **5+ AI product features** using prompt engineering and agentic design patterns, compressing delivery and validation cycles by **50%** through cross-functional collaboration.

CFM, RCOEM

May 2025 - Jul 2025

Machine Learning Research Intern

Nagpur, India

- Developed a **clinical sleep-disorder prediction system** trained on **1,000+ patient health records** with biometric feature engineering (blood pressure, sleep duration, stress indices), achieving **87% prediction accuracy**.
- Optimized end-to-end **ML data pipelines** (Python, Scikit-learn, Pandas) via cross-validation and rigorous model evaluation, cutting development time by **30%** and boosting validation accuracy by **25%**.

PROJECTS

Autonomous Executive Email Copilot

[GitHub](#)

Multi-Agent RAG, LangChain, FastAPI, SciPy, React

- Built a **multi-agent Agentic AI framework** with a deterministic RL-style simulation environment for benchmarking autonomous email agents across classification, prioritization, and full action execution.
- Implemented **4 policy modes** (baseline heuristic, perturbation stress-test, LLM-driven, multi-agent hybrid) scored against bounded, numerically stable grading metrics.
- Delivered **RAG-powered context-aware prioritization** via vector retrieval and semantic search, achieving an estimated **60% reduction** in inbox triage time.
- Instrumented a full ops stack - FastAPI telemetry pipelines, approval workflows, episode replay, and structured audit trails - enabling reproducible agent evaluation.

DealSentry - Intelligent Proposal & Compliance Guard

[GitHub](#)

Azure OpenAI, RAG, React, Express.js, Prisma, PostgreSQL

- Shipped an **enterprise AI compliance platform** (Azure OpenAI, RAG) automating proposal validation with AI-driven risk scoring and DOCX/PDF ingestion, reducing manual review effort by **70%**.
- Devised a **contract lifecycle tracking engine** with a rule-based compliance engine, role-based access control, approval routing, SLA tracking, and audit logging.
- Integrated **Salesforce, HubSpot, and Gmail APIs** plus Puppeteer PDF generation and automated document parsing to eliminate manual handoffs across sales and compliance.

SentinelOps - AI-Powered Incident Detection Platform

[GitHub](#)

FastAPI, IsolationForest, statsmodels, NetworkX, React, Supabase

- Orchestrated an **autonomous incident prediction and prevention platform** processing **10,000+ synthetic telemetry events** with IsolationForest anomaly detection and dependency-graph modeling.
- Automated **multi-agent root-cause analysis**, business-impact estimation, and human-approved remediation, reducing estimated MTTD by **80%** versus manual monitoring.
- Configured **Slack and PagerDuty alerting** with observability pipelines, structured audit logging, and optional Kubernetes / webhook hooks for real infrastructure.

LifePulse - Health Analytics & Prediction Platform

[GitHub](#)

Scikit-learn, XGBoost, ONNX, Pandas, Flask, Gunicorn

- Trained **4 Random Forest and XGBoost models** for heart disease (**87.2%**), sleep disorder (**78.3%**), migraine (**82.0%**), and personalized health scoring (**0.82 R²**).
- Constructed **feature pipelines** over BMI, blood pressure, cholesterol, and lifestyle factors with risk stratification (Python, Scikit-learn, Pandas, NumPy).
- Deployed models with **ONNX / onnxruntime** for offline inference and Git LFS model versioning on Render (Flask, Gunicorn), with a USDA FoodData-integrated nutrition tracker.

Shadow GTM - Multi-Tenant SaaS Intelligence Platform

[GitHub](#)

Next.js, React 19, TypeScript, Supabase, Stripe, Recharts

- Launched a **multi-tenant SaaS platform** with workspaces, role-based access, autonomous scheduling, and **Stripe billing**.
- Exposed a **public REST API** with outbound webhooks and multi-channel alerts (email, Slack, webhooks), plus live web-data scanning, diffing, and source-grounded signal extraction.
- Established **analytics dashboards** (Recharts), schema validation (Zod), error monitoring (Sentry), and automated testing (Vitest) for reliable releases.

EDUCATION

Shri Ramdeobaba College of Engineering and Management (RCOEM)	2022 - 2026
B.Tech, Electronics & Communication Engineering	GPA: 8.90 / 10
Minor, Artificial Intelligence & Machine Learning	GPA: 9.67 / 10

CERTIFICATIONS & AWARDS

- AWS Certified Cloud Practitioner - Amazon Web Services (Oct 2025)
- Hackathon Finalist - Paytm x Sarvam x Logitech AI National Hackathon
- 2nd Place - ByteSize Sage AI National Hackathon

AREAS OF INTEREST

Applied AI & LLM Product Engineering, Agentic & Multi-Agent Systems, Retrieval-Augmented Generation, MLOps & Model Deployment, Full-Stack Development, AI Infrastructure & Observability.