

Raghav Sampath

(646) 647-5633 | rs4760@columbia.edu | [linkedin.com/in/raghavsampath/](https://www.linkedin.com/in/raghavsampath/) | github.com/thehyperpineapple | s-raghav.vercel.app | US Citizen

EDUCATION

Columbia University

Master of Science in Data Science

Relevant Coursework: Machine Learning, Agentic AI, Statistical Inference and Modeling

New York City, NY

Dec 2026

Vellore Institute of Technology

Bachelor of Technology in Electronics and Communication Engineering

Relevant Coursework: Probability and Statistics, Artificial Intelligence and Machine Learning, Linear Algebra

Vellore, India

Jun 2025

EXPERIENCE

Acceedo

Software Engineering Intern | JavaScript, MongoDB, Python, RestAPIs

- Built a **real-time data pipeline** for IoT telemetry, enabling faster data availability for analytics and decision-making.
- Optimized **time-series data storage** in MongoDB, improving data quality and enabling efficient exploratory analysis
- Partnered with operations to translate business requirements into **data-driven pricing logic**, improving automation & consistency

Coimbatore, India

Jan 2025 – Apr 2025

Research Assistant - Patent Filed (Application No. 202541122085)

Real-time Foundry Monitoring System under Dr. Christina Josephine Malathi A

- Developed an **IoT-based data collection system** for environmental monitoring, generating high-frequency datasets for analysis
- Built a **real-time dashboard** to visualize streaming data and monitor key metrics for operational insights
- Implemented secure APIs and data pipelines to ensure **reliable and consistent data availability for analytics use cases**

Vellore, India

Jan 2025 – Apr 2025

Quadrasystems.net

Machine Learning Engineer Intern | Python, AWS, Anomaly Detection, NLP, PyTorch

- Developed a **vulnerability prediction model** using PyTorch, improving accuracy by **12%** through advanced NLP feature engineering
- Fine-tuned **DistilBERT** using transfer learning, reducing training time by **~30%** while maintaining model performance
- Deployed a **real-time ML inference service**, enabling actionable predictions and communicating results to non-technical stakeholders

Coimbatore, India

Jun 2024 – Aug 2024

PROJECTS

Morpheus | Python, FastAPI, Next.js 15, TypeScript, Vertex AI, GCP Cloud Run, Tailwind, shadcn/ui

Adversarial Multi-Agent Strategy Simulation

- Built a **holonic multi-agent** engine where personas debate strategic scenarios in structured rounds, surfacing second-order effects.
- Engineered a **FastAPI + Vertex AI** backend with **SSE streaming**, retry logic, and structured JSON logging; keeping schemas in sync.
- Shipped a **Next.js 15 + Tailwind** frontend with **XY Flow** scenario trees and **Recharts** risk matrices on a single **Cloud Run** service.

Apr 2026

Better Call Scout | Python, Google ADK, RAG, ChromaDB, GCP Cloud Run

Multi-Agent AI Startup Scout

- Built **multi-agent LLM system** (Google ADK) with parallel signal collectors & multi-perspective hypotheses evals with confidence scoring
- Designed a **dynamic RAG system (ChromaDB)** enabling critic-agent grounding in prior startup data and filtering low-signal inputs
- Deployed on **Google Cloud Run**, aggregating VC, developer, & journalist agent outputs into structured memos with visualizations

Apr 2026

OrderFlow | Python, NLP, LLM, FastAPI, PostgreSQL

AI-Powered B2B Order Processing Platform

- Built an **NLP-driven extraction system** to convert unstructured text into structured data, achieving **95%+ accuracy** across **10K+ SKUs**
- Developed an **anomaly detection service** to auto-flag suspicious or unusually large orders for review, reducing manual effort.
- Engineered a **FastAPI** backend with **PostgreSQL JSONB** schema storing dynamic line items, analytics layer with revenue trends.

Feb 2026

SKILLS

Languages & Tools: Python, SQL, R, C++, Git, Docker, JavaScript, TypeScript, AWS (EC2, S3), GCP (BigQuery, Vertex AI), MongoDB.

Frameworks: PyTorch, TensorFlow, FastAPI, Google ADK, Scikit-learn, Spark, Hadoop, XGBoost, LightGBM, Hugging Face Transformers, Statsmodels, Pandas, NumPy, Matplotlib, Seaborn, OpenCV, LangChain, LangGraph, Node.js, Express, Claude Code, Codex, Cursor.

Techniques: Machine Learning, Deep Learning, Retrieval Augmented Generation (RAG), Agentic Workflows, Natural Language Processing (NLP), Time-Series Analysis, Statistical Modeling, A/B Testing, Large Language Models (LLMs), MLOps, Anomaly Detection.

POSITIONS OF RESPONSIBILITY

Voting Member of the **Data Science Institute Student Council** at **Columbia University**

Elected from **40,000 students** as **Executive Member, VIT Student Council**

Jan 2026 – Present

Sep 2023 – May 2025