

Arushi Gupta

MA | arushigupt.23@gmail.com | linkedin.com/in/arushig05 | github.com/i4rushy | US Permanent Resident

WORK EXPERIENCE

ReVenture AI

Lead Agentic AI Workflow Engineer

Dedham, MA

May 2025 – August 2025

- Architected and shipped a **production AI agent** for autonomous outbound sales, designing the agent's reasoning, prompts, and tool-call architecture across **OpenAI GPT-4o**, ElevenLabs TTS, and GHY — replacing manual outreach and delivering real-time customer impact across multiple Fortune-adjacent client deployments.
- Designed an **agentic decision-routing layer** where the LLM dynamically selects one of four downstream tool calls (close sale, book meeting, tag CRM, trigger nurture sequence) based on real-time intent classification, demonstrating end-to-end ownership of how the agent thinks and acts.
- Built **end-to-end backend integrations** by authoring against REST and GraphQL-style APIs across GHY, Notion, Mailchimp, and Calendly, implementing retry logic, dead-letter queuing, audit logging, and GDPR/Do-Not-Call compliance for production-grade scalability.
- Partnered directly with **client stakeholders and sales teams** to translate business challenges into agentic solutions, leading technical scoping conversations and iterating on deployments that improved outreach throughput by 35%.

Purdue School of Electrical & Computer Engineering

Data Structures Teaching Assistant

West Lafayette, IN

August 2025 – December 2025

- Led technical discussions with 200+ students, breaking down complex algorithmic and OOP concepts into clear, accessible explanations — sharpening the consultative communication skills required to lead high-level technical conversations with non-technical stakeholders.
- Reviewed and gave engineering-rigor feedback on student code, evaluating correctness, efficiency, and design — the same rigor applied to evaluating AI-generated outputs in production agent workflows.

Outlier AI

AI RLHF Training Engineer

Remote

May 2024 – December 2024

- Evaluated and ranked LLM outputs across coding, reasoning, and NLP tasks with engineering rigor, providing structured human feedback that drove alignment fine-tuning cycles for frontier large language models.
- Stress-tested model responses for factual accuracy, logical consistency, and prompt-following behavior — the same evaluation discipline required to ship reliable agentic systems into customer production environments.
- Documented findings in actionable reports consumed by cross-functional ML research teams, supporting rapid iteration on model quality.

EDUCATION

Purdue University

Bachelor of Science in Computer Engineering (Concentration in AI & ML)

West Lafayette, IN

May 2026

Relevant Coursework: Machine Learning, Natural Language Processing, Data Mining & Analysis, Object-Oriented Programming, Statistical Methods, Linear Algebra, Design & Analysis of Algorithms, Database Systems

SKILLS

Languages: Python, JavaScript/TypeScript, Rust, SQL, C/C++, R

AI & Agentic Systems: LLM Prompt Engineering, Agent Architecture & Tool Calling, Agentic Workflow Design, LLM Fine-Tuning & RLHF Evaluation, Retrieval-Augmented Generation (RAG), GraphRAG, Transformer Models (Hugging Face), Text Classification

Web & APIs: React, REST APIs, GraphQL, FastAPI, Actix-web, OOP Design, JSON Schema Validation

Real-Time & Data Infra: ETL/ELT Pipeline Design, Real-Time Data Streaming, Apache Kafka, Pub/Sub, Concurrency Control

Cloud & MLOps: Google Cloud Platform (GCP), Cloud Run, Cloud Scheduler, GitHub Actions (CI/CD), Docker

Tools & Platforms: OpenAI API, Anthropic Claude, Cursor, n8n, ElevenLabs API, Git/GitHub, PostgreSQL, Neo4j, Qdrant

PROJECTS

GraphRAG Pipeline – Production Agentic Retrieval System

August 2025

Engineered a **production-grade agentic GraphRAG pipeline in Rust** where an LLM agent autonomously extracts structured knowledge graphs from unstructured documents and dynamically routes queries across hybrid retrieval modes for downstream reasoning. Designed the full agent architecture — semantic chunking, **LLM-based entity/relation extraction via structured prompt engineering and tool calling**, Neo4j graph construction with Louvain community detection, and Qdrant vector indexing. Implemented **local and global search modes** enabling both fact retrieval and thematic summarization across large corpora — solving reasoning tasks that defeat vanilla RAG. Deployed a **REST API with caching, retry logic, and concurrency control** via Actix-web; benchmarked latency (P50/P95) and answer quality across 4 question categories against a Python baseline, demonstrating Rust's throughput advantage under concurrent load.

Real-Time News Intelligence Pipeline – Cloud AI & Real-Time Infra

January 2026

Architected a **real-time cloud pipeline on GCP** ingesting live news streams via Pub/Sub and Cloud Scheduler, orchestrating parallel **LLM-based summarization, sentiment classification, and topic tagging** across 500+ articles/day through Cloud Functions — with retry logic, dead-letter queuing, and schema validation for full pipeline observability. Indexed enriched embeddings into **Qdrant** for semantic search with sub-second retrieval latency, and deployed a **FastAPI query service on Cloud Run** exposing REST endpoints for semantic search, sentiment trends, and topic clustering. Automated end-to-end deployment via **GitHub Actions CI/CD**, reducing time-to-insight from hours to seconds.