

ADITYA SURESH

Sydney, NSW | P: +61 468 442 565 | hello@adityasuresh.com | Portfolio | LinkedIn

PROFILE

AI Engineer with 3+ years of experience building scalable data science platforms, predictive analytics workflows, and production-grade automation across AWS and Azure cloud environments, bridging applied ML and software engineering. Proficient in Python, SQL, applied ML, MLOps, and CI/CD, with experience turning complex operational data into reliable pipelines, model-driven workflows, and intelligent automation systems. Recently specialized in Generative AI through a Master's in Artificial Intelligence at UNSW, building practical Retrieval-Augmented Generation (RAG) systems, agentic workflows, and LLM-integrated applications using LangGraph, LangChain, MCP, FastAPI, and modern cloud tooling.

EDUCATION

UNIVERSITY OF NEW SOUTH WALES (UNSW)

Master of Information Technology (Artificial Intelligence), WAM: **81 (Excellence)**

Sydney, AU

Sep 2023 - Aug 2025

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Bachelor of Technology in Computer Science & Engineering

Trivandrum, IN

Aug 2016 - Jul 2020

WORK EXPERIENCE

INFOSYS

Data Scientist

Technopark, IN

Nov 2020 - Aug 2023

- Selected among the **top 1%** of Infosys engineers (Power Programmer track); designed and delivered analytics and ML-driven solutions for multinational enterprise clients, including a **Fortune 500 rail client**, while building robust Python/SQL data processing workflows across Snowflake and AWS for high-volume operational datasets from multiple heterogeneous sources.
- Developed and deployed predictive and optimization models using Python (pandas, NumPy, scikit-learn) and R for routing efficiency, asset utilization, and anomaly detection; implemented model training and batch inference pipelines on AWS (S3, EC2, Lambda, SageMaker, Athena) to enable production-grade decision support systems.
- Engineered end-to-end ML workflows including feature engineering, model evaluation, hyperparameter tuning, and validation; utilized MLflow for experiment tracking and model versioning, improving reproducibility and reducing model comparison effort by **~30%** across environments.
- Operationalized ML systems using Docker and Kubernetes with CI/CD workflows (Jenkins, GitHub Actions) to automate testing, deployment, and monitoring; improved system reliability and reduced manual validation effort by **~40%** while maintaining SLA-driven reliability.

PROJECTS

Multi-Agentic Flight Booking Assistant | LLMs (Gemini), LiteLLM, LangGraph, MCP, FastAPI

Jan 2026 - Present

- Built an agent-driven flight booking system using Gemini LLMs, LiteLLM fallback routing, LangGraph, MCP-style tool orchestration, FastAPI, and Typescript; implemented structured intent extraction, IATA/date validation, API-backed offer search, authentication, admin controls, and saved trip workflows, reducing incomplete searches by **~50%** in simulated testing.
- Architected a production-ready price-tracking workflow with Turso/libSQL persistence, Celery/RabbitMQ scheduled scans, run history, offer normalization, quota controls, Langfuse/OpenTelemetry/Grafana observability, user feedback capture, Github CI quality gates, and LLM evaluations for reliable public rollout.

AI-Powered HR Policy Assistant | RAG, LangChain, Ollama, ChromaDB, RAGAS

May 2026 - Present

- Building a RAG-based HR policy assistant for natural-language access to unstructured company policy documents; implemented semantic retrieval, employee metadata filtering, source-backed responses, and context-aware answer generation for role, location, employment type, and policy applicability.
- Architecting the ingestion and governance layer with PDF parsing, chunking, embeddings, ChromaDB indexing, admin-controlled policy updates, access controls, query logging, RAGAS-style evaluation, and fallback guardrails to support auditable enterprise HR knowledge workflows.

SKILLS

Programming & Data Science: Python, SQL, pandas, NumPy, scikit-learn, TensorFlow, PyTorch, NLP, R, React
Generative AI & Agentic Systems: LLMs, RAG, AI Agents, LangChain, LangGraph, LlamaIndex, HuggingFace, OpenAI APIs, Anthropic APIs, LiteLLM, Ollama, AWS Bedrock, Pydantic AI, DSPy, MCP, CrewAI, AutoGen, Prompt Engineering
Retrieval, Search & Vector Systems: ChromaDB, FAISS, Pinecone, Weaviate, Elasticsearch, Hybrid Search, Semantic Search, BM25 + Dense Retrieval, Custom Embeddings
Cloud, MLOps & Deployment: AWS, Azure, GCP, Vertex AI, Azure Synapse, Snowflake, MLflow, Weights & Biases, Kubeflow, LangSmith, Langfuse, RAGAS, TruLens, LLM-as-a-Judge, Docker, Kubernetes, FastAPI, Redis, CI/CD Pipelines, Guardrails