

SWARNAVA DUTTA

swarnava.dev@gmail.com | +91-7478893810 | [linkedin.com/in/swarnava-dutta](https://www.linkedin.com/in/swarnava-dutta)

Senior Generative AI Engineer with 6 years of experience architecting and shipping production-grade Agentic AI systems, multi-agent frameworks, and enterprise-scale RAG pipelines. Delivered AI copilots and autonomous agent workflows using LangChain, LangGraph, and Model Context Protocol (MCP), with LLM fine-tuning (LoRA/PEFT), prompt engineering, and RAGAS-driven evaluation, backed by FastAPI REST APIs and JavaScript/TypeScript (ReactJS). Experienced with Azure AI Foundry, Copilot Studio, and Azure AI Search, plus AWS, deploying via Docker and Azure Kubernetes Service (AKS).

EXPERIENCE

Senior Generative AI Engineer | Technical Lead

Dec 2025 - Present

TEKsystems Global Services Pvt Limited

- Spearheaded a multi-agent conversational system using **LangGraph** orchestration, unifying governance and development requests through context-aware questioning, cutting intake-to-ticket turnaround from 2+ days to under 10 minutes.
- Compressed **P95** pipeline latency by 85% through parallelized field derivation across concurrent Redis-backed workers, while delivering rule-based risk classification and automated dual-platform ticketing with **human-in-the-loop** approval checkpoints.
- Architected an end-to-end agentic learning platform using **Azure AI Foundry** and semantic retrieval from **Azure AI Search** to generate adaptive quizzes and detect learner skill gaps, supporting 6k+ monthly sessions and improving detection accuracy by 37% under continuous **RAGAS** evaluation.

Senior Generative AI Engineer | Technical Lead

Jul 2025 - Dec 2025

Syren Technologies Private Limited

- Led optimisation and scaling of a supply chain copilot with **LangGraph**, **Pinecone**, and **FastAPI**, serving 10K+ daily queries, reduced latency by 63%, and ensuring 99.9% uptime with resilient ingestion and monitoring pipelines.
- Fine-tuned **Mistral-7B** with **LoRA** on supply chain datasets, boosting contextual recall by 34% and improving contract/invoice automation efficiency evaluated using **BLEU/ROUGE-L** (generation quality) and retrieval relevance metrics (recall@k).
- Designed and managed context injection, **prompt chaining**, and secure orchestration using **MCP** in Gen AI workflows, improving multi-turn response accuracy by 28% and reducing context errors by 42%.

Senior Generative AI Engineer | Technical Lead

Jan 2024 - Jul 2025

RWS Moravia India Private Limited

- Managed the company's first enterprise AI agent and production-grade **multilingual RAG** platform, cutting manual knowledge lookup by 43% while instituting rigorous **BLEU** and **NDCG** benchmarks to drive measurable model performance.

Junior Data Scientist

Jul 2022 - Dec 2023

BinaryERP Software Private Limited

- Orchestrated pipelines to support **embedding** generation and lightweight text classification, enabling native **semantic search** prototypes adopted by 13 B2B users, while improving dataset usability by 26%.

Associate Data Engineer

Jan 2021 - Jul 2022

Capgemini Technology Services India Limited

- Implemented delta-based ingestion into **Azure Data Lake** for a global banking client, gathering and refining data requirements, reducing nightly batch runtime by 95% and failures by 61%, ensuring compliance with strict financial **SLAs**.

SKILLS

- **Generative AI:** LangGraph, LangChain, LlamaIndex, Model Context Protocol (MCP), Agentic AI, Multi-Agent Systems, Hugging Face Transformers, LoRA/PEFT, Prompt Engineering, BERT, A2A Protocol, Anthropic Claude, Human-in-the-Loop (HITL)
- **Cloud & AI Platforms:** Azure AI Foundry, Azure OpenAI, Copilot Studio, Azure AI Search, Azure Functions, Key Vault, AWS
- **Programming & APIs:** Python, JavaScript/TypeScript, REST API design (FastAPI), React, SQL, PyTorch, TensorFlow, PySpark, Pandas, NumPy
- **RAG & Vector Search:** Pinecone, FAISS, ChromaDB, Pgvector, Hybrid Search, Semantic Chunking
- **Evaluation:** RAGAS, LangSmith, Langfuse, BLEU, ROUGE, Recall@k, Precision@k, NDCG, F1
- **LLMOps & DevOps:** Docker, Azure Kubernetes Service, Azure Container Registry, Redis, Git, CI/CD, Azure DevOps, MLflow
- **Data Engineering:** Azure Databricks, ADF, Delta Lake, Unity Catalog, Kafka, Data Modelling

PROJECTS

Agentic AI Ticket Automation System

- Built an intent-triage architecture with conditional routing requests across specialized agent workflows, transforming untracked email intake into a single auditable, session-persistent channel, with accuracy of 98%

MedTech Sales Multi Agent Platform

- Designed a persona-adaptive agentic sales enablement system leveraging k-means clustering and Redis caching for low-latency content retrieval, AI-generated simulations that improved rep assessment scores by 30% through performance optimisation.

Hybrid Agentic Enterprise Chatbot

- Led development of a multi-agent conversational AI combining RAG and SQL agents over Databricks Delta Tables via Genie MCP, handling 10K+ monthly queries and cutting analyst turnaround time by 42% through intelligent hybrid query routing

Multi Agent Recruiter Assist Tool

- Directed development of an in-house recruiter-assist AI tool using LlamaIndex, OpenAI embeddings, Pinecone, and FastAPI, evaluated with F1-score to validate extraction accuracy, reducing screening workload by 33% on 1k+ resumes.

Swiss Legal RAG platform

- Extended a Swiss client's voice-to-text platform with a RAG chatbot for contract drafting and clause summarization, enabling 45% faster legal document turnaround while preserving GDPR compliance.

HR and Finance Data Integrated Chatbot

- Developed a domain-specific AI assistant with Databricks ETL, LangChain, and Azure OpenAI, integrated into Cognitive Search, cutting routine HR queries by 40% and saving 200+ staff hours monthly.

EDUCATION

Maulana Abul Kalam Azad University of Technology
Bachelor of Technology, Computer Science and Engineering

Aug 2016 - Jun 2020

CERTIFICATIONS

- **Microsoft:** [AZ-900 Azure Fundamentals](#)
- **Microsoft:** [DP-900 Azure Data Fundamentals](#)
- **Databricks:** [Databricks Fundamentals](#)
- **Databricks:** [Generative AI Fundamentals](#)
- **Databricks:** [Azure Databricks Platform Architect](#)
- **Apache:** [Apache Airflow Fundamentals](#)