

# SWARNAVA DUTTA

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

Senior Generative AI Engineer with 5 years of experience in AI/ML. Spearheaded enterprise adoption of Agentic AI copilots, Model Context Protocol (MCP) workflows, and multimodal RAG systems for large-scale automation. Delivered production-grade copilots using LangChain, LangGraph, and LlamaIndex with LoRA/PEFT fine-tuning, deployed on Azure Kubernetes Service (AKS). Experienced with Azure AI Foundry, Copilot Studio, and Azure AI Search, along with vector databases (Pinecone, FAISS, Chroma). Strong foundations in PySpark, Databricks, Azure Data Factory, and MLOps (Docker, FastAPI, CI/CD).

## EXPERIENCE

### Senior Generative AI Engineer | Technical Lead

Dec 2025 - Present

*TEKsystems Global Services Pvt Limited*

- Led development of a voice enabled multi-agent analytics copilot using **LangChain**, **LangGraph** and **MCP**, over **Snowflake semantic views**, reducing analyst effort by 67% through automated SQL generation while supporting 500+ monthly business queries.
- Designed an MCP-driven RAG platform using **vector embeddings** and **semantic retrieval** to ingest learning content, generate quizzes, and detect learner skill gaps, supporting 6k+ monthly sessions and improving detection accuracy by 37%.
- Engineered an end-to-end agentic learning engine with confidence scoring and adaptive content regeneration, integrating **Redis** for data caching to reduce redundant DB calls improving concept retention by 42% and reducing redundant study time by 56%.

### 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*

- Led 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

- GenAI and LLMs:** RAG, Agentic AI, Model Context Protocol (MCP), LangGraph, LangChain, LlamaIndex, Hugging Face Transformers, OpenAI, Azure OpenAI, LoRA/PEFT, Prompt Engineering, BERT, T5, AI Agents
- Azure AI Stack:** Azure AI Foundry, Copilot Studio, Azure AI Search, PromptFlow
- Evaluation Metrics:** BLEU, ROUGE, Recall@k, NDCG, F1, RAGAS, LangSmith
- Vector Databases and Search:** Pinecone, FAISS, ChromaDB, Pgvector, Hybrid Search, Semantic Search
- LLMOps:** Git, Docker, Kubernetes, FastAPI, Redis, CI/CD, Airflow, Monitoring and Logging
- Programming:** Python, PyTorch, TensorFlow, SQL, Pandas, NumPy
- Data Engineering:** Azure Databricks, ADF, Delta Lake, Unity Catalog, Kafka, Azure DevOps, PySpark, Data Modeling

## PROJECTS

### MedTech Sales Enablement Platform

- Designed an agentic sales enablement system with persona-adaptive RAG and MCP, backed by Redis caching to serve frequent content lookups with low latency, AI-generated simulations, skill gap detection, and closed-loop performance optimization.

### Voice-Enabled Agentic Data Copilot

- Engineered a voice-enabled multi-agent AI copilot using LangGraph and MCP over Snowflake semantic views to translate natural language queries into SQL, delivering insights in under 10 seconds across 8 enterprise datasets.

### Hybrid Agentic Enterprise Chatbot

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

## Agentic AI Content Generator

- Orchestrated an Agentic AI system using LangGraph to autonomously research and generate LinkedIn posts, integrating Kafka for real-time content stream ingestion, reducing content creation time by 84% through parallel processing and boosting engagement via quality scoring and hook A/B testing.

## Talksmith - RAG Chatbot

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

## Swiss Legal RAG chatbot

- Developed a legal-tech chatbot for Swiss client, extending the dictation platform with a RAG chatbot for contract drafting and clause summarization. Built a secure ingestion and semantic chunking process, 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**

**Aug 2016 - Jun 2020**

*Bachelor of Technology, Computer Science and Engineering*

## 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](#)