

RAGHAV MEHTA

8770014305 | raghavmehta1801@gmail.com | linkedin.com/in/raghav-mehta | github.com/Raghavmehta01

TECHNICAL SKILLS

Languages: Python, SQL

Web Tools: HTML, CSS, XML, Django

Tools & Platforms: MS Fabric, Azure Data Factory, AI Foundry

Other Concepts: Data Science, GenAI, Machine Learning, ETL, Power BI

Libraries: Scikit-learn, NumPy, Pandas, Matplotlib

EXPERIENCE

Data Engineering Intern

Orange Data Tech

June 2025 – Present

Indore, MP

- Worked on client-facing data engineering projects involving the development and optimization of large-scale ETL pipelines in MS Fabric and Snowflake using Azure Functions, Python Notebooks, SQL, Spark, and AI Foundry. Implemented Medallion architecture (Bronze–Silver–Gold) to improve data quality and analytics readiness. Contributed to AI/ML and Generative AI initiatives using Llama and Mistral LLMs, containerizing solutions with Docker and deploying them on Azure.
- **Logistics Supply & Order Processing ETL Pipeline (MS Fabric):** Designed and deployed ingestion and transformation pipelines using Fabric Data Pipelines, Spark Dataflows, Stored Procedures, REST APIs, and Lakehouse layers. Automated supplier-SKU mapping, order ingestion, inventory reconciliation, and validation workflows with incremental refresh and error-handling mechanisms.
- **PEO Domain ETL Pipeline (Snowflake, Notebooks, APIs):** Built Snowflake-based ETL pipelines for HR, payroll, and employee datasets using Python API ingestion, Snowflake Notebooks, and SQL Views. Implemented Bronze/Silver/Gold layers with automated cleansing, schema harmonization, SCD handling, and curated analytical models optimized for reporting teams.
- **AI & GenAI Integrations (AI Foundry, LLMs):** Developed LLM-powered automation workflows for data classification and document summarization. Built prompt-engineering modules and deployed Dockerized GenAI microservices integrated with Snowflake and Fabric pipelines, including vector store creation and embedding workflows.

Data Analyst Intern

Brainy Ladder

March 2024 – June 2024

Indore, MP

- Performed ETL operations to integrate and transform raw datasets into structured formats, followed by creating interactive Power BI dashboards and reports to visualize key metrics.
- Collaborated with cross-functional teams to gather requirements, analyze trends and anomalies, and deliver actionable insights for decision-making.

PROJECTS

RAGify – The Smart Chat Companion ([Link](#))

Python, LangChain, ChromaDB, LLM, Docker, Azure

- Developed a Retrieval-Augmented Generation (RAG) chatbot using LangChain and ChromaDB for document retrieval, embedding, and contextual response generation.
- Integrated multi-LLM support (Azure-AI, OpenAI GPT, Google Gemini) and built a modular Python backend with a lightweight HTML/JS interface.
- Containerized the application with Docker and deployed on Azure Web Apps for scalable vector-based operations with persistent Chroma storage.

BloodNet – A Leukemia Detection Model ([Link](#))

Deep Learning, CNN, Django

- Designed and trained a CNN model to detect leukemia from microscopic blood cell images with preprocessing, normalization, and augmentation pipelines.
- Developed a Django-based web application for real-time detection and visualization.

Electoral Dynamics in Madhya Pradesh

Python, Power BI

- Transformed raw election data and performed EDA to extract trends at constituency and state levels.
- Created interactive Power BI dashboards to visualize electoral outcomes and performance metrics.

EDUCATION

Lovely Professional University

Master of Computer Applications

Aug 2023 – June 2025

Jalandhar, Punjab