

Transforming Railway IT with Scalable Data and AI Solutions

नई दिल्ली



NEW DELHI

نئی دہلی

प्रवेश
ENTRY

टिकट काउंटर
TICKET COUNTERS

RAILWAY
ASSISTANT

CUSTOMER SUCCESS STORY

ABOUT THE CLIENT



INDUSTRY: PUBLIC SECTOR - RAILWAYS

The client is a leading Public Sector Undertaking under the Ministry of Railways, responsible for developing and managing critical IT infrastructure for the Indian Railways. They drive digital transformation and innovation across the railway ecosystem.



BUSINESS CHALLENGES



On-premises environment limitations affecting scalability and performance.



The need for an advanced Data Lake to handle large datasets.



Implementation of MLOps use cases with Databricks for AI initiatives.



Lack of scalability, flexibility, and computing power for modern data workloads.



TECHNOLOGY USED

01



MICROSOFT AZURE CLOUD

02



AZURE DATABRICKS

03



AZURE DATA FACTORY (ADF)

04



MEDALLION ARCHITECTURE
(BRONZE, SILVER, GOLD)

SOLUTIONS OFFERED



Provisioned Azure Cloud infrastructure and Azure Databricks to enable distributed data processing and ML-driven workloads.



Built a scalable and secure Data Lake architecture for efficient data management.



Implemented end-to-end MLOps workflows covering model development, training, and deployment.



Designed robust data pipelines using Azure Data Factory with secure Site-to-Site VPN integration for seamless on-premises database ingestion.



Adopted Medallion Architecture for structured data processing (Bronze: raw data, Silver: cleansed/enriched data, Gold: refined/aggregated datasets).



BUSINESS IMPACTS



Provisioned Azure Cloud infrastructure and Azure Databricks for distributed data processing and ML workloads.



Established a scalable and secure Data Lake architecture.



Implemented end-to-end MLOps workflows for model development, training, and deployment.



Designed data pipelines using Azure Data Factory with a secure Site-to-Site VPN for ingestion from on-premises databases.



Adopted Medallion Architecture for structured data processing (Bronze: raw data, Silver: cleansed/enriched data, Gold: refined/aggregated datasets).

