

# Chetan Hiwale

Data Engineer

+91-8483015532 | Hiwalechetan73@gmail.com | Pune, INDIA | Experience: 3 Years 1 Months

Data Engineer with expertise in optimizing data models and SQL queries, achieving a 30% reduction in execution time. Skilled in building CI/CD pipelines using Azure DevOps to streamline Microsoft Fabric deployments. Experienced in automating data ingestion with Data Factory and debugging PySpark jobs to enhance efficiency. Led POCs and training sessions to drive team adoption of new technologies. Designed scalable pipelines processing data and developed automated migration solutions, reducing manual effort by 80%.

## SKILLS

- **Programming & Query Languages:** Python, Scala, SQL, Kusto Query Language (KQL).
- **Data Platforms & Big Data Processing:** Apache Spark / PySpark, Spark Streaming, Hadoop, Databricks, AWS EMR, AWS Glue, Microsoft Fabric, ADLS Gen2, Data Modelling, ETL, Power BI.
- **Streaming & Real-Time Data:** Apache Kafka, Azure Event Hub.
- **Orchestration, Workflow & DevOps:** Apache Airflow, Azure Data Factory (ADF), Azure DevOps, Docker, LangChain, CrewAI.

## WORK EXPERIENCE

Data Engineer | LTI Mindtree – Pune, India

Sept 2023– Present

### Project: Azure COE

1. **Optimized data model** by normalizing tables and **enhancing SQL query performance**, resulting in a 40% reduction in execution time for metadata ingestion system.
2. Developed **automated migration solutions** for Fabric data pipelines and notebooks using **Airflow**, reducing manual effort
3. Build SSIS to Pyspark based converter from scratch which uses parsing and **AI agents** to create pyspark code for Databricks and Microsoft Fabric.
4. Designed and developed a **spark-based Metadata ingestion framework which covers Data Quality, validation and ingestion.**
5. Conducted **code reviews and mentoring sessions** promoting Spark and Microsoft Fabric best practice, performance optimization across team and organization.

### Project: PSE

1. **Developed and optimized end to end ETL process and scalable Data pipelines in medallion architecture** to process structured and unstructured datasets and ensured efficient downstream consumption
2. **Collaborated closely with clients, stakeholders and other cross functional teams** to understand the requirements as per the SOW.
3. **Implemented data ingestion and transformation from APIs, AWS S3, Athena in Fabric Lakehouse** with best practices (caching, partitioning).
4. **Developed reusable scripts and workflows** to streamline environment creation, reducing manual effort by 40%.
5. **Debugged monitored and optimized performance PySpark jobs** via Spark UI to identify, resolve performance bottlenecks, improving job efficiency.
6. Ensured the fabric items and other dependencies **CI/CD and DevOps optimized.**

## **Associate Trainee | LTI-Mindtree – Pune, India**

Jan 2023-August 2023

1. Created automated python script for Fabric workloads using Fabric Rest APIs.
2. Got hands on experience on SSIS tool for data movement and transformation and Power BI reporting.

## **Project: Safety Sphere (Microsoft Fabric Hackathon): 1 Months**

1. Developed a real-time location tracking system to monitor user activities and deliver instant alerts, enhancing user safety and engagement.
2. Designed and implemented advanced geofencing solutions to dynamically classify safe and unsafe zones, improving risk management accuracy.
3. Transformed and streamed real-time data utilizing Azure Event Hub and Kusto Query Language (KQL) to generate actionable, intuitive reports for stakeholders.

## **CERTIFICATIONS AND AWARDS:**

1. Star Performer award, Problem solving and Accountability Badge.
2. Microsoft Certified: Fabric Analytics Engineer Associate
3. AWS Associate Data Engineer
4. Databricks Associate Data Engineer
5. Microsoft Certified: Fabric Data Engineer Associate

## **EDUCATION:**

2023 B.Tech/B.E. | Chemical Vishwakarma Institute of Technology, Pune