Hemanth Chowdisetti

Contact: +91 **9014980496**

Mail: hemanthkumarchowdisetti999@gmail.com

PROFILE SUMMARY

Results-driven **Azure Data Engineer** with 3.6 years of experience in **data engineering, business intelligence, and analytics**. Skilled in designing and optimizing **data pipelines, ETL/ELT processes, and cloud-based data solutions**. Strong expertise in **Azure Data Factory, Azure Data Lake Storage, Azure Synapse Analytics (SQL Data Warehouse), and Databricks with Apache Spark** for large-scale data integration and transformation.

PROFESSSIONAL EXPERIENCE

- Azure Data Engineer with 3.6 years of experience in **building and managing data pipelines, ETL workflows**, and **cloud-based analytics** solutions.
- Worked on Azure Services **Azure Gen2lake, Azure Key Vault, Azure Data Bricks, Azure Data Factory, SQL Server.**
- Experience in building **Azure Data Bricks (ADB) PySpark-Python, Spark-SQL** Notebooks to perform data transformations.
- Implemented **Delta Lake on Azure Data Lake Storage,** enabling reliable **ACID transactions** and improving data quality for downstream analytics. Expert in reading and writing data in **Azure Data Bricks.**
- Experience in big data components such as **Spark SQL**, **Data-frames** and **Hive** database implementation using **Azure Data bricks**
- Good hands-on experience implementing end to end **Data Migration** includes both Data ingestion using **Azure Data Factory**.
- Having hands on maintaining **Delta Lake** and maintains data in different zone Bronze, Silver and Gold. Implemented Meta Data driven **ADF pipelines** to process the incremental Data Load.
- Worked in **ADF components** Linked services, Data sets, Integration runtime, Pipelines, Activities and Triggers.
- Implemented pipelines using **ADF activities** like Lookup activity, for each activity, copy activity, notebook activity, Stored procedure activity, set variables, get metadata activity, delete activity, conditional statements.
- Worked with **Agile methodology** (sprint planning, estimation and tracking of deliverables to Ensure delivery of quality.
- Good Built automated testing frameworks within **CI/CD pipelines** for validating data quality, schema consistency, and pipeline performance during the deployment lifecycle.

Technical Skills

Database   : SQL
ETL   : Azure Data factory
Azure Services: Azure Databricks, Azure Data factory, Azure Key Vault,
                Delta Lake, AdlsGen2, Azure Synapse, Git & Git Hub, CI/CD (Azure DevOps), Hadoop cluster,
                Map reduce, YARN, Airflow, and Power BI.
Languages   : Python, SQL

EXPERIENCE SUMMARY

Working as a **Data Engineer** in **Technomold IT Solutions Pvt Ltd** from **February - 2022** to **Till Date**

MAJOR PROJECTS UNDERTAKEN:

Client: M&T Bank   **Duration:** Sep -2024 to Till Date

Project **M&T Bank – Enterprise Data Platform Modernization**

Role: Consultant - Data Engineer

Environment: Azure Data Factory, Databricks, Azure Synapse Analytics, ADLS Gen2, Event Hubs/Kafka, SQL, Power BI, Git, Azure DevOps

Description: M&T Bank wanted to modernize its legacy on-premises data warehouse and ETL pipelines to a scalable, cloud-native data platform on Azure. The goal was to handle real-time data ingestion, faster reporting, regulatory compliance, and customer analytics.

Responsibilities:

- I have maintained and optimized an Azure Data Factory–based data orchestration framework to **automate ETL/ELT workflows, ensure data reliability, and support analytics workloads**
- Oversaw data from on-premise SFTP servers to **Synapse**.
- Created metadata-driven generic **ADF** pipelines.
- Performed data transformation and testing at different data layers using **Databricks** notebooks.
- Conducted data quality checks using **PySpark** & **Spark SQL** vi a **Databricks**.
- Handled PII data and Logic App workflows for pipeline triggers, ensuring smooth data orchestration from on-premise to Azure environments. Also, responsible for monitoring and supporting PROD pipelines and addressing issues when they arise.
- Automated deployment of **SQL data pipelines** across environments, reducing release cycle time by **40%**.
- Integrated **real-time data ingestion pipelines** using **Event Hubs** for customer behaviour tracking and sales transactions.
- Integrated **unit testing & code quality checks** into **CI/CD pipeline** for reliable data pipeline delivery.
- Implemented **CI/CD pipelines in Azure DevOp**s integrated with **GitHub**, automating deployments of **Azure Data Factory pipelines and Databricks notebooks.**

Client: **MUFG Bank, Ltd. (Mitsubishi UFJ Financial Group)**   **Duration:** March -2022 to Aug -2024

Project: MUFG Bank – Enterprise Data Platform

Role: **Consultant – Azure Data Engineer**

Environment: Azure Data Factory, Databricks, Azure Synapse Analytics, Data Lake Storage Gen2, Event Hubs, SQL, Git & Git Hub, CI/CD (Azure DevOps).

Domain: Finance & Banking

Description: MUFG Bank is one of the world's largest financial institutions with 772 domestic and 76 overseas branches. The goal of the project was to build a secure, scalable, and cloud-based data platform to support regulatory compliance, risk management, and business insights by migrating legacy systems to Azure cloud.

Roles & Responsibilities:

- Designed and developed **ETL/ELT** pipelines in **Azure Data Factory (ADF)** to ingest data from on-premises databases, flat files, and APIs into **Azure Data Lake Storage (ADLS).**
- Applied complex data transformations using **Data Flow activities in ADF, including lookups, aggregations, joins,** and conditional splits **to implement business logic.**
- Optimized fact and dimension table loading **in Azure Synapse Analytics**, **leveraging PolyBase**, **partitionin**g, and indexing for performance tuning.
- Configured parameterized pipelines and **reusable datasets**, **linked services**, and triggers for dynamic and scalable data integration.
- Orchestrated workflows with event **triggers, control flow activities** (If, Until, Switch), and integrated with Azure Functions for advanced logic.
- Implemented incremental data loading using watermarking and **Change Data Capture** (CDC) patterns.
- Developed and managed Delta Lake tables on **Azure Databricks** for incremental batch loads and **ACID transactions**.
- Performed performance tuning using partition pruning, cache, broadcast joins, and pushdown optimization in Databricks and Synapse.
- Secured data pipelines with **Azure Key Vault**, **RBAC policies**, and **network security** groups (NSGs).
- Collaborated with cross-functional teams for **Data migration, testing, and production support**, ensuring smooth cutovers and root cause analysis of incidents.
- Managed version control for data engineering projects using **Git and GitHub,** ensuring seamless collaboration and code quality across distributed teams.
- Created and maintained feature branches, handled pull requests, and performed code reviews to enforce best practices in data pipeline development.
- Implemented **CI/CD pipelines in Azure DevOps integrated with GitHub**, automating deployments of **Azure Data Factory pipelines** and **Databricks notebooks.**

ACADEMIC CREDENTIALS

Completed **Degree in BSc** from **Sr Venkateswara University, Tirupati** in Dec - **2021**