Mohammad Hakeem

Data engineer

✉ mohammadabdulhakeem555@gmail.com   📞 : 8106625168   📍   Hyderabad

in   Hakeem-mohammad-abdul

PROFILE

Experienced data engineer with 3 years of experience having strong background in designing and implementation scalable data architectures. Proficient in ETL processes and optimization. Skilled in utilizing technologies such as Apache Spark, Python, SQL, Airflow and Azure data technologies.

WORK EXPERIENCE

Infosys Limited (Data Engineer)
Digital Specialist Engineer

➤ **TECHOPS**
   Tech Stack – Pyspark, Python, Spark SQL, Yarn, AWS Airflow, GitHub Copilot-AI, AWS S3 datalake ,Minio datalake.

   - Created an application to track technician activities, including the time taken to install new products at customer homes and the number of installations completed by technician on each day.
   - Developed complex and optimized batch data pipelines to process   **60 million records daily** as per business requirements.
   - Achieved a **60% reduction in execution time** through performance optimization, enhancing overall efficiency.

➤ **INTERACTION DATAMART**
   Tech Stack – Databricks, Pyspark, Spark SQL, Python, Minio data lake.

   - Build data pipeline to pull the data from multiple digital channels(web,XA,IVR,chat) into DataMart.
   - To increase the effectiveness of each channel by measuring the time it takes to resolve each issue.
   - It helped to Reduce the frequency of creating complex ad hoc queries to pull data from disparate sources.
   - Designed the pipeline to efficiently handle **GB's** of data on a daily basis.

➤ **E- COMMERCE SALES**

   Tech Stack – Azure Databricks, Azure Data Factory (ADF), Azure Data Lake Storage (ADLS), Azure Key Vault, Pyspark, Spark SQL, Python.

   - Built a data pipeline using **Medallion Architecture** for organized and efficient data processing.
   - Extracted data from diverse sources such as Azure SQL and REST APIs using Azure Data Factory, and orchestrated Databricks notebooks for efficient data processing.
   - GitHub project link : https://github.com/Dataengg-Ninja/Azure_Project

➢ **ALTHLETE METRICS SYSTEM**
  Tech Stack – *Azure Databricks, Azure Data Factory (ADF), Azure Data Lake Storage (ADLS), Azure Key Vault, Pyspark, Python.*

- Built a data pipeline to **migrate on-premises SQL Server data to Azure Cloud.**
- Leveraged ADF to ingest data from on-premises SQL Server into Azure Data Lake Storage, orchestrating the entire data pipeline seamlessly within ADF.
- Leveraged Databricks for advanced data transformation and processing.
- GitHub project link : https://github.com/Dataengg-Ninja/onprem_azure

CERTIFICATIONS AND ACHIEVEMENTS

- SQL(Basic & Intermediate) HackerRank
- Infosys Insta Award
- Infosys Certified Spark Professional
- Infosys Certified Python Programmer

EDUCATION

Bachelor of Technology (Computer Science of Engineering)

Vaageswari College of Engineering   *07/2017 - 07/2021*

SKILLS

- Python
- SQL
- PySpark
- AWS Airflow
- Azure Databricks,
- Azure Data Factory (ADF)
- Azure Data Lake Storage
- AWS S3