

MOBEEN MEHDI

Islamabad, 44000 Pakistan
+923324971110 - mobeenmehdi@outlook.com

WEBSITES, PORTFOLIOS, PROFILES

- [linkedin.com/in/mobeenmehdi/](https://www.linkedin.com/in/mobeenmehdi/)

PROFESSIONAL SUMMARY

Accomplished data engineer with a strong background in solution architecture, specializing in designing and optimizing robust data pipelines and ETL processes. Expertise in data warehousing, data modeling, and Big Data processing frameworks like Apache Spark. Skilled in SQL and relational databases, with a focus on performance tuning and maintaining high data quality standards. Proven track record in cloud infrastructure management, adept at implementing DevOps methodologies to streamline deployment and enhance scalability. Demonstrated leadership in promoting data governance and integrating real-time analytics to drive actionable insights and strategic decision-making.

SKILLS

- | | |
|------------------------|-----------------------------------|
| • ETL Development | • Data Governance |
| • Data Warehousing | • Real-time Analytics |
| • Data Modeling | • Data Quality Assurance |
| • Data Pipeline Design | • Relational databases |
| • Data Migration | • RDBMS |
| • Big Data Processing | • Serverless Computing |
| • Spark Framework | • Cloud Infrastructure Management |
| • Performance Tuning | • Linux System Administration |
| • SQL Expertise | • DevOps Methodologies |

WORK HISTORY

06/2023 to Current

Data Platform Engineer

TERADATA - GDC PAKISTAN – Remote

- **Data Platform Engineer | Target Australia** : Designed and developed a scalable data platform, improving data processing speed by 35% and reducing storage costs by 25%.
- Integrated data from over 20 internal and external sources, enhancing data quality and consistency across the organization.
- Implemented automated CI/CD pipelines, reducing deployment time by 50% and increasing platform reliability.
- Optimized data pipelines, resulting in a 40% increase in data throughput and a 20% reduction in latency.
- Enhanced inventory management and demand forecasting through real-time data analytics, leading to a 15% reduction in stockouts and a 10% increase in sales.

01/2023 to Current

DATA ENGINEER

TERADATA - GDC PAKISTAN – Remote

- **DIGI Cell Malaysia** : Designed and implemented scalable data streaming pipelines with AWS Kinesis Firehose reducing data processing times by 40% and enabling real-time analytics.
- Utilized Kinesis Firehose to deliver streaming data to S3, Redshift, and Elasticsearch, improving data accessibility and reducing data retrieval times by 30%.
- Optimized migration processes and streaming pipelines, improving client workload efficiency by 50% and enhancing overall system performance.
- Collaborated with cross-functional teams to meet data requirements and provided comprehensive post-migration support and documentation.

07/2021 to 01/2023 CONSULTANT - DATA ENGINEER

KEYRUS DELIVERY CENTER - PAKISTAN – Remote

- **Data Engineer | Essilor luxottica : Data Migration:** Utilize Azure Data Factory to extract and migrate data from on-prem Postgres to Azure Data Lake, improving migration efficiency by 40%.
- **Data Lake Architecture:** Implement a three-tier architecture (L1, L2, L3) in the data lake to enhance data retrieval efficiency by 30%.
- **Data Integration:** Integrate various data sources into a unified format using ADF and Databricks, reducing data silos by 50%.
- **Data Transformation:** Apply data integrity and transformation using ADF Mapping Dataflows/Databricks, increasing data accuracy by 25%.
- **Error Handling:** Develop Logic Apps for robust error handling in ADF data pipelines.
- **Incremental Load & ARM Templates:** Implement incremental load and snapshot data ingestion strategies, and develop ARM templates for consistent deployment across environments.
- **SAP BO Integration:** Build SAP BusinessObjects (BO) views from Databricks views, enabling semantic layer data access and KPI development.

12/2021 to 12/2022 Data Engineer

KEYRUS DELIVERY CENTER - PAKISTAN

- **Data Engineer - Borealis** : Developed Metadata-Driven Ingestion Framework Automated data ingestion from JDBC and file sources, reducing manual effort by 40% using Apache Airflow and Azure Big Data services.
- Customized ingestion pipelines using dynamic dags for flexibility and scalability, optimizing data workflows.
- Managed seamless migration, minimizing downtime by 50% during transition.
- Implemented efficient data integration pipelines, enhancing storage efficiency by 25% using Mapping Dataflow.
- Created templates for consistent pipeline deployment across environments.
- Developed DBT code for dynamic snapshot-based transformations, ensuring data reliability and reporting accuracy.

04/2022 to 06/2022 Data Bricks Unity Catalog POC -Data Engineer

Keyrus – Lahore, Punjab, Pakistan

- **Unity Catalog POC - Johnson & Johnson** : Deployed AWS Databricks

Lakehouse orchestrated deployment using Cloud Formation on AWS achieving a 30% reduction in deployment time and costs.

- Successfully conducted a POC demonstrating a 40% improvement in data lineage clarity and governance efficiency.
- Developed robust ETL pipelines resulting in a 25% increase in data processing speed and accuracy within Databricks environment.
- Implemented effective data governance strategies using Databricks Unity Catalog, enhancing data integrity and compliance by 35%.
- Centrally managed and governed all data assets across different Databricks workspaces, ensuring 50% faster access and utilization.

05/2019 to 01/2020 **IOS DEVELOPER**

DEVSTOP PVT

- Data Pipeline - IOS Developer May 2019 - Jan2020
- Responsibilities: Client - Valet Tracker
- Worked on the Valet Tracker app, which is basically used for car parking and helps drivers find parking easily
- Handled 3 phases of this app, "Client Side", "Driver Side" and "Admin Side"

EDUCATION

08/2021

Master of Science: Software Engineering

Fast Nucs - Islamabad

Thesis Summary:

- Leveraged the Cleveland dataset to conduct a comprehensive analysis of feature selection techniques in cardiovascular heart disease prediction.
- Applied ensemble learning techniques, specifically bagging models, to enhance predictive accuracy and robustness of cardiovascular risk prediction models.
- Employed machine learning algorithms to evaluate and compare the effectiveness of feature subsets in predicting cardiovascular heart diseases, using metrics like AUC-ROC and calibration plots.
- Investigated the interpretability and clinical relevance of selected features to refine predictive models and improve understanding of cardiovascular risk factors.
- Advanced research contributing to the optimization of healthcare diagnostics through rigorous validation and integration of bagging models in predictive analytics for cardiovascular health assessment.
- Proposed implications for future research, emphasizing the potential of ensemble learning in personalized medicine and healthcare decision support systems.

- Professional Development Studies: Big Data, 2020

10/2017

Bachelor of Science: Software Engineering

University of Management And Technology Pakistan - Lahore, Punjab, Pakistan

- **Final Year Project**
- Collaborated in a team of two to develop a mobile application prototype similar to Food Panda, focusing on food ordering and delivery services.
- Designed and implemented an intuitive user interface and interactive user

- Developed backend server-side logic using modern frameworks to manage user accounts, restaurant listings, and order processing.
- Integrated real-time order tracking and notifications to enhance customer satisfaction and operational efficiency for delivery partners.
- Conducted usability testing and gathered feedback as a team to iteratively improve application features, ensuring optimal performance and user engagement.
- Presented joint findings and recommendations for future enhancements, emphasizing scalability and market competitiveness in the food delivery sector.

Title: Data Engineer

- AWS Data Analytics - Specialty, Keyrus - 2022
- Databricks Lakehouse Fundamentals, keyrus - 2022
- Vantage Associate 2.3, Teradata - 2023
- Vantage Administration , Teradata - 2024

Urdu: Native language

English: C2

Proficient