NAVYA TEJA VANKA

AZURE DATA ENGINEER | navyateja169@gmail.com | LinkedIn | +91 9182410129

Azure Data Engineer with 3.11 years of hands-on experience in designing, developing and maintaining data solutions using Azure technologies. Proficient in Azure Data Factory, Azure Databricks, Azure Synapse and other Azure Services.

Strong experience in data virtualization using Denodo and developing BI solutions with MicroStrategy. Proven ability to deliver high-quality, scalable solutions.

TECHNICAL SKILLS:

- **Data Engineering:** PySpark, Azure Databricks, Azure Data Factory (ADF), ETL/ELT Development
- **Data Virtualization:** Denodo
- **Programming & Query Languages:** Python, SQL, Spark SQL
- **Cloud Platforms:** Microsoft Azure
- **BI & Reporting Tools:** MicroStrategy
- **DevOps (Foundational):** Git, Azure DevOps
- **Databases & Warehouses:** SQL Server, Oracle, SSMS
- **Data Architecture:** Data Lake, Data Warehouse, Delta Lake
- **Job Scheduling Tools:** Control-M, Stonebranch

PROFESSIONAL SUMMARY:

Data Engineer **Accenture – Hyderabad,India**

Client: Qurate Retail Group Jan 2024 – Present

- Built and optimized data ingestion and data transformation frameworks in **Azure Databricks** using **PySpark**, Delta Lake, Data Frames and Medallion Architecture (Bronze/Silver/Gold).
- Buit data pipelines in Databricks with spark performance to speed up processing and reduce resource usage for 2 TB+ datasets
- Worked on data migration from legacy systems (Oracle, Teradata, Netezza) to Cloud Common Data Model **(CDM)** and **Synapse** to standardize enterprise reporting.
- Managed CI/CD deployments in Azure DevOps for Databricks notebooks, ADF Pipelines and configuration-based release management
- Created data quality validation rules, schema checks, and exception handling frameworks for improving reliability and auditability of ETL jobs.
- Performed large-scale historical backfill (2TB+) in Delta Lake optimizing file sizes, Z-order, vacuum, optimize, enabling better downstream analytics performance.
- Implemented metadata-driven ETL and orchestration using Azure Data Factory with dependence handling, parallel execution, triggers and alerts.

- Worked in Agile/Scrum environment participating in sprint planning, grooming, and daily standups.
- Scheduled and monitored pipelines using **Stonebranch**, ensuring SLA compliance and timely data delivery.

Client: VISA Jan 2022-Jan 2024

- Created and optimized **Denodo** base, derived, join, union, and interface views to deliver virtualized datasets for enterprise reporting.
- Developed **PySpark** transformations in Azure Databricks for data masking, deduplication, aggregation, anomaly detection and transaction-level processing.
- Built Delta Lake pipelines implementing time travel, versioning, MERGE (upserts), streaming + batch ingestion for analytics and BI reporting
- Implemented data quality validation and error logging frameworks integrated within Databricks notebooks.
- Implemented **event-driven ADF triggers** and monitoring alerts for SLA adherence, data freshness, and failure notifications.
- Automated metadata-driven workflows combining ADF + Databricks for scalable ingestion and transformation.
- Wrote optimized PySpark DataFrame code and advanced SQL (window functions, CTEs, aggregations) for large-scale transformations
- Ensured uninterrupted data flow by configuring **pipeline retries, alerts, and monitoring dashboards**.
- Used MicroStrategy Developer, Web, Object Manager, and Command Manager to design, build, and deploy BI solutions.

CERTIFICATIONS:

- AZ-900: Azure Fundamentals
- AWS Certified Cloud Practitioner
- DP-900 - Microsoft Azure Data Fundamentals

EDUCATION:

Bachelor of Technology, Gayatri Vidya Parishad College of Engineering, Visakhapatnam
Aug 2021 | CGPA: 7.91