Adwet Naithani

Dehradun,Uttarakhand,India-248001

📞 8755682089 ✉ adwetnaithani@gmail.com in adwet adwet007

Summary

Databricks Certified Data Engineer Associate with over 1 year of hands-on experience, including a 5-month internship and 8 months in a full-time Data Engineer role. Strong foundation in data analytics from a prior 6-month internship, with proven skills in building scalable data pipelines, optimizing ETL workflows, and working with cloud and big data technologies

Certification

- Databricks Certified Data Engineer Associate
- Astronomer Certification for Apache Airflow 2 Fundamentals
- Databricks Accredited Lakehouse Fundamentals
- Google Data Analytics Professional Certificate
- Microsoft Power BI Data Analyst by Microsoft
- IBM Data Analyst Professional Certificate
- Data Science with Databricks for Data Analyst

Work Experience

Digivatelabs   **December 2024 – Present**
Data Engineer

- Proven expertise in migrating legacy systems (Cloudera, Redshift, Vertica) to Databricks, enhancing performance and reducing latency.
- Built scalable data pipelines using Apache Spark and Databricks, leveraging Medallion Architecture.
- Reduced job runtime by 20% by optimizing data processing workflows.

Client – Payment Gateway

- Spearheaded the migration of 30TB of critical user data from Amazon S3 to Databricks Unity Catalog, ensuring schema consistency, secure access control, and data integrity across development, staging, and production environments.
- Transformed cumbersome Amazon QuickSight SQL queries, optimizing for faster data retrieval and diminishing cloud costs by 30%, enabling data scientists to focus on strategic projects.
- Automated the parameterization of 10+ SQL queries in Databricks, ensuring consistent and scalable execution across 4,000+ instances, resulting in standardized data retrieval processes for data analysts.

Client – E-commerce marketplace

- Migrated over 10,000 tables from Vertica to Databricks, accounting for 30% of the total data migration effort, while ensuring schema consistency, data accuracy, and minimal system downtime.
- Boosted query performance by 35% during data migration by applying efficient partitioning on year, month, and day columns, significantly improving data retrieval speed and minimizing scan time.
- Implemented transformation logic supporting complex nested types (List, Map or nested struct) with prefix handling and conditional rules from JSON metadata.
- Reduced data migration time by 40% through the implementation of parallel processing techniques for high-volume table ingestion, improving overall project efficiency and scalability.

Digivatelabs   **July 2024-November 2024**
Data Engineer Intern

Client – Entertainment and Online Ticketing-Poc

- Applied query tuning strategies including partitioning, indexing, and caching—improving speed by 30%.
- Reduced resource utilization and execution time by 20% through PySpark UDF tuning.
- Conducted performance profiling of ETL workloads to identify and fix bottlenecks.

Technical Skills

Programming	Python, R, SQL (MySQL, PostgreSQL,Apache Spark SQL)
Data Processing	PySpark, Pandas, Delta Lake
Big Data / Cloud	Databricks, Data Lakehouse, AWS(S3,Redshift,RDS), Azure
Data Analysis	Power BI,Pandas, Matplotlib, Seaborn, Data Wrangling
Automation	Databricks Workflow,Apache Airflow,Version control(Git),DLT
ETL / Governance	Batch and Real time ingestion ,Data Modeling, Data Governance
Format	Parquet,JSON,Orc,Avro,CSV

Education

Graphic Era Hill University
Master of Technology   *Dehradun, Uttarakhand*
University Of Petroleum and Energy Studies
Bachelor of Technology   *Dehradun, Uttarakhand*