

MD Ali

+1(603) 233-9080 ♦ Charlotte, NC mdalietl200@[gmail.com](mailto:mdalietl200@gmail.com) ♦ [linkedin.com/in/md-ali-a2b38118a](https://www.linkedin.com/in/md-ali-a2b38118a) ♦ <https://ajanbiali200.github.io/>

OBJECTIVE

Data Engineer with 5+ years of experience in Big Data, Cloud, and Data Warehousing, seeking full-time Data Engineering or Big Data roles where I can leverage Spark, Snowflake, AWS, and Azure experience, contributing to Agile teams through strong collaboration and expertise in end-to-end data solutions.

EDUCATION

Master of Science in Computer Science, Rivier University

Sep 2022 – May 2024

SKILLS

Big Data: Hadoop, Spark (Scala/PySpark), Hive, Kafka, Sqoop, NiFi, Oozie, File Formats (CSV, XML, JSON, Parquet, Delta), Processing (Batch Processing).

Cloud Platforms: AWS (EC2, S3), Azure (ADF, Databricks, DevOps), Databricks (Medallion Architecture).

Data Warehousing: Snowflake, SQL Server, Oracle, MySQL, DB2, Data Modeling (Facts/Dimensions, Star/Snowflake Schema).

Languages: Python, SQL, Shell Scripting, C#.

Tools: Power BI, DBT, Git, Jira, Control-M, Splunk.

Methodologies/Practices: Agile Methodology, Data Validation, Data Quality Checks.

EXPERIENCE

Data Engineer | Duke Energy – Charlotte, NC |

Jun 2024 – Present

- Built scalable ETL data pipelines using Spark (Scala) on Databricks, implementing batch processing for large-scale data ingestion and transformation. Optimized data loading into Snowflake, handling various file formats (e.g., Parquet, Delta).
- Implemented performance optimization and tuning using Spark for enhanced data processing efficiency.
- Worked on AWS infrastructure (EC2, S3) and deployed Kafka clusters with secure integrations. Actively participated in Agile methodology and fostered strong team collaboration for project success.
- Developed Power BI dashboards using data from HDFS and Snowflake, incorporating rigorous data validation and quality checks to ensure accuracy and reliability.

Data Engineer | CVS Health – Remote |

Jun 2023 – May 2024

- Developed robust ETL workflows using PySpark on Azure Databricks into Snowflake, implementing incremental and full load strategies to efficiently populate Snowflake data warehouses.
- Automated data pipelines using Azure Data Factory and managed them with Azure DevOps, adhering to Agile methodologies and promoting team collaboration throughout the development lifecycle.
- Built Power BI dashboards, implementing role-based security using Snowflake, and ensuring data integrity through comprehensive data validation and quality checks.
- Led migration efforts from Hive/Pig to Spark, resulting in significant performance optimization and tuning for data processing.

ETL Developer | UST Global – India |

Jun 2020 – Sep 2022

- Migrated data from Netezza to Snowflake, utilizing ETL processes and building DBT models and ADF pipelines. Managed data with a focus on incremental and full load mechanisms.
- Integrated Power BI for interactive reporting, managing complex data workflows on Azure. Applied data validation and quality checks to maintain high data standards.
- Automated data ingestion and transformation using Python and SQL scripting, processing various file formats (e.g., CSV, XML, JSON) and implementing efficient batch processing routines.

Software Developer Intern | UST Global – India |

Jun 2019 – May 2020

- Developed web apps using C#, React, and MongoDB.
- Deployed services in Azure cloud with version control and CI/CD.