Bommana Hareesh Reddy

Software Engineer

✉ bommana.hareesh@gmail.com 📞 8310254005 📍 Bangalore, Karnataka

Professional Summary

- Azure Data Engineer with 3.8+ years of experience in designing, developing, and managing scalable **ETL pipelines**, **data lakes**, and analytics solutions using **Microsoft Azure** ecosystem including **Azure Data Factory, Azure Databricks, Azure SQL, and Azure Data Lake Storage Gen2.**
- Proficient in **Spark, PySpark, SQL,** and Python for big data processing. Adept in **Agile delivery, version control with Gi**t, and enterprise-level data migration projects.
- Strong hands-on experience with **PySpark** for big data processing, transformations, and aggregations, alongside experience working with Hadoop ecosystems.
- Optimized **Spark jobs** for better performance using techniques like **partitioning, caching, and broadcast joins**.
- Built **ETL pipelines** using **PySpark and Delta Lake** to manage both **structured and unstructured data efficiently**.
- Designed and improved **complex SQL queries, indexing strategies, and execution plans** for high-performance reporting.
- Experience in **data modeling** and working with **Azure SQL Database** for efficient storage and retrieval.
- Developed **end-to-end ETL pipelines** in **Azure Data Factory**, integrating data from multiple sources.
- Automated data workflows using **Triggers, Scheduling, and Parameterized Pipelines** for seamless data movement.
- Organizing data into the **Medallion Architecture (Bronze, Silver, Gold layers).**
- Managed **structured and unstructured data** using **Azure Data Lake Storage Gen 2** and **Azure Blob Storage.**
- Worked in **Agile environments**, following **Scrum methodologies** for project planning, sprint execution, and continuous delivery.

Skills

Languages:
Python, SQL, PySpark

Cloud Services:
Azure Data Factory, Azure SQL, Azure Blob Storage, Azure Data Lake Gen2, Azure Key Vault, Azure Monitor

Methodologies:
Agile, Scrum

Big Data Technologies:
Apache Spark, Delta Lake, Azure Databricks

Tools:
Git, GitHub, Jira, ServiceNow

Operating Systems:
Windows, Linux

Professional Experience

HCLTech – Software Engineer	**01/2025 – Present**
IBM – Software Developer	**09/2021 – 01/2025**

PROJECTS

PROJECT : 3
Client: Optum Healthcare.   **01/2025 - Present**
Domain: HealthCare
Tools: Azure Data Factory, Azure Databricks, PySpark , Azure Delta Lake, SQL
Roles and Responsibilities:

- Managed robust data pipelines with **Azure Data Factory**, reducing **ETL processing** time and enabling real-time analytics that enhanced decision-making capabilities.

- Engineered and optimized data pipelines using **Azure Data Factory,** significantly reducing data processing time and improving overall data accuracy
- Leveraged **Azure Databricks** for advanced data processing and machine learning integration.
- Developed and implemented scalable Spark clusters within **Azure Databricks** to handle large volumes of data efficiently.
- Developed and enhanced Spark jobs and notebooks in **Azure Databricks** to perform data transformations, aggregations, and analysis.
- executed data lake solutions using **Azure Data Lake Storage Gen2**, providing scalable storage for structured and unstructured data.
- Integrated Azure Data Lake Storage with Azure services such as Azure Data Factory and Azure Databricks to enable seamless data ingestion, scalable data processing, and advanced analytics workflows.
- Established version control practices using **Git** or Azure Repos to manage the codebase and configuration files, ensuring versioning consistency and collaboration among team members.
- Actively participated in **Agile ceremonies** including sprint planning, daily standups, sprint reviews and retrospectives to ensure alignment with project goals and prirorities

PROJECT : 2
Client: Lumen Technologies.   **03/ 2023 – 01/ 2025**
Domain: E-commerce..
Environments: Azure Data Factory, Azure Databricks, PySpark ,Azure Delta Lake, SQL
Roles and Responsibilities:
- Developed scalable **ADF pipelines**, improving data throughput and reducing failures.
- Assisted in integrating **Azure Databricks** Notebooks with **ADF** for advanced data processing during the migration process.
- Created diverse **ADF pipelines** specifically tailored to address the requirements.
- Configured **Azure cloud services**, including **Data Lake Storage** and **Azure SQL Database**, optimizing them for efficient data migration and storage.
- Implemented email notification features using Azure Logic Apps for real-time monitoring of the data migration progress.
- Developed scripts and automation tools to facilitate data **extraction, transformation, and loading processes**.
- Conducted data validation and quality checks to ensure the accuracy and integrity of migrated data.
- Collaborated closely with senior team members to understand the intricacies of the data migration process and implement effective solutions.

PROJECT : 1
Client: AT & T.   **09/ 2021 – 03/ 2023**
Domain: E-commerce.
Environments: Azure Data Factory, Azure Databricks, PySpark ,Azure Delta Lake, SQL
Roles and Responsibilities:
- Designing and developing **ETL pipelines** using **Azure Data Factory** to
- Implementing error handling mechanisms and monitoring solutions within **ADF** pipelines to ensure smooth and reliable data processing
- Developing and optimizing **Spark jobs** and notebooks in **Azure Databricks** to perform data transformations, aggregations, and analysis
- Architecting and implementing data lake solutions using **Azure Data Lake Storage Gen2**, providing scalable storage for structured and unstructured data
- Integrating **Azure Data Lake Storage** with **Azure services** like **Azure Data Factory** and **Azure Databricks** to enable seamless data ingestion, processing, and analytics workflows.
- Planned and executed scalable Spark clusters within Azure Databricks to handle large volumes of data efficiently.

EDUCATION

Bachelor of Engineering
Dr. k v Subbareddy Institute of Technology   **2016 - 2020**