Lohith Reddy Meruva

DATA ENGINEER

Maryland, USA | Mobile: (203) 910-7427 | Email: lohithmde@gmail.com

SUMMARY

- **3+ years of experience in data engineering** with a focus on designing, implementing, and optimizing data solutions.
- Proficient in programming languages such as **Python, SQL, and R,** with expertise in utilizing essential packages like **NumPy, Pandas, and Matplotlib** for data manipulation and analysis.
- Demonstrated proficiency in leveraging cloud technologies, particularly **Microsoft Azure and Google Cloud Technology**. Extensive experience with **Azure Data Factory, Azure Databricks, BLOB storage, and Azure SQL Server** for scalable and collaborative data solutions.
- In-depth knowledge of big data technologies, including **Hadoop, MapReduce, HDFS, Sqoop, Hive, NIFI, Kafka, and Apache Spark**. Applied these technologies to optimize data processing and analytics.
- Utilized **Git and GitHub** for version control, ensuring collaborative and organized development processes.
- Worked closely with healthcare professionals, data scientists, and business analysts to understand data requirements and deliver effective solutions.

SKILLS

Languages:	Python, SQL, R
Packages:	NumPy, Pandas, Matplotlib, TensorFlow
Database Management:	MySQL, PostgreSQL, MongoDB, Oracle
Cloud Technologies:	Microsoft Azure, Azure Data Factory, Azure Databricks, BLOB, Azure SQL Server, GCP
Big Data Technologies:	Hadoop, MapReduce, HDFS, Sqoop, Hive, NIFI, Kafka, Apache Spark
IDEs:	Visual Studio Code, PyCharm, Juypter Notebook
Visualization Tools:	Tableau, Power BI, Quick Sight, Looker
Other Technical Skills:	Machine Learning, Data Management, Marketing Analytics, Information Technology Strategy, Selenium, Jira, Digital Innovation, ETL/ELT Process Innovation & Management, Big Data Technology, Microsoft SQL Server, SSIS, SSRS, SSAS, Kubernetes, Informatica, Natural language processing
Version Control Tool:	Git, GitHub, SVN, Bitbucket

EXPERIENCE

Optum, MD **Jan 2023 – Present**

Data Engineer

- Designed and implemented systems to integrate and consolidate healthcare data from diverse sources, such as electronic health records (EHRs), claims data, and patient registries.
- Developed infrastructure for real-time processing of patient vitals using Azure Stream Analytics, reducing data latency by 20%.
- Used Apache Spark and Hadoop to optimize healthcare data processing, resulting in a remarkable 40% reduction in processing times.
- Developed Extract, Transform, Load (ETL) pipelines to process, clean, and transform raw healthcare data into a structured format for analysis.
- Built and maintained a healthcare data warehouse on Azure SQL Data Warehouse, enabling efficient storage and retrieval of 100+ TB of healthcare data.
- Implemented Apache Kafka for real-time data streaming, enabling near real-time healthcare analytics and faster decision-making for healthcare providers.
- Implemented Azure Data Lake Storage for scalable data storage and Azure Databricks for collaborative data analytics, enhancing the overall data ecosystem.
- Worked closely with healthcare professionals and data scientists to understand requirements, ensuring alignment with business goals and objectives.

Birla Soft, India **Jul 2019 – Dec 2021**
Data Engineer

- Implemented Azure Data Factory (ADF) extensively for ingesting data from different source systems like relational and unstructured data to meet business functional requirements.
- Designed and implemented ETL processes using tools like Informatica, achieving a 25% reduction in data integration time.
- Designed and developed Batch processing and real-time processing solutions using ADF, Databricks clusters and stream Analytics
- Developed automation scripts to streamline data processes, reducing manual intervention and increasing operational efficiency by 25%.
- Leveraged Oracle PL/SQL and MS SQL server development skills to create and maintain database structures, stored procedures, and functions.
- Configured Spark Streaming to receive real time data from the Apache Kafka and store the stream data to HDFS using Python.
- Worked with complex SQL views, Stored Procedures, Triggers, and packages in large databases from various servers.
- Collaborated with data scientists to deploy machine learning models, improving predictive analytics accuracy by 18%.
- Worked closely with stakeholders, data scientists, and business analysts to understand data requirements and deliver effective solutions.
- Conducted data analysis to support business decision-making by extracting, cleansing, and manipulating data from various sources.
- Created data visualizations to communicate complex data sets in an easily understandable format for business users.

EDUCATION

Master of Professional Studies in Data Science - University of Maryland, Baltimore County, Baltimore, MD
Bachelor of Technology in Computer Science and Engineering - Lovely Professional University, Punjab, India