Lohit Marla | lohitmarla@gmail.com | GitHub | www.linkedin.com/in/lohitmarla/
https://lohitmarla-uconn.github.io/portfolio/   +1 (860) 931-8028   Hartford, Connecticut

Education

University of Connecticut, Storrs, CT   December 2024
Master's in Data Science, 4.0 / 4.0 GPA
GITAM University, Vizag, India   August 2020
Bachelor of Technology, Computer Science and Engineering, 8.82 / 10 GPA

Skills

Programming: Python, Scala, Pyspark, SQL
Data Engineering: Airflow, GitHub, ETL, Data Warehousing, Data Mining, Data Cleaning, Data Modeling
Cloud Technologies: AWS, Terraform, Snowflake Fundamentals, Databricks
Data Analysis and Visualization: Pandas, NumPy, Tableau, GitHub, Generative Artificial Intelligence Fundamentals
Soft Skills: Communication, Team Collaboration, Attention to Detail, Time Management, Adaptability

Work Experience

Data Engineer, Novus Platform Vizag, India   September 2019 - July 2023
ETL and Data Pipelines

- **Designed and Developed Data Solutions**: Spearheaded the design, development, and maintenance of data models and ETL/ELT pipelines using Airflow and Pyspark, ensuring timely and high-quality data availability across data platforms and warehouses.
- **Data Ingestion and Integration**: Engineered and implemented Incremental Data Ingestion Glue Pipelines for Xifin and Lims, integrating daily transactional data and ensuring seamless data flow.
- **Data Transformation and Quality**: Crafted Apache Spark-based data comparison systems, automating data quality controls and data reporting in Tableau. Developed synthetic data using Python and Synthetic Data Vault (SDV) to enable effective testing and validation by downstream teams.
- **Optimization and Performance Tuning**: Played a key role in performance tuning of Spark jobs by reviewing Spark architecture, optimizing execution efficiency, successfully reducing job runtime from 8 hours to less than 50 minutes.
- **Cost Optimization**: Implemented a cost-saving initiative to eliminate the dependency on MySQL for data storage, transitioning to S3 data storage solution, resulting in an 80% cost reduction while maintaining seamless data access and integrity.
- **API Data Processing Workflow**: Built API data processing workflows using Serverless AWS technologies (Appflow, Lambda, SQS, SNS, DynamoDB) and terraform, reducing manual effort and streamlining data processing.
- **Data Analysis and Troubleshooting**: Provided technical support, troubleshooting, and resolving bugs through root cause analysis using SQL. Conducted performance tuning and optimization on batch data workflows, saving costs by up to 55%.
- **Leadership and Collaboration**: Led cross-functional teams to gather business requirements, leverage data processing technologies, and construct data extraction and transformation pipelines for Google Campaign Manager, aligning with business logic and compliance.
- **Subject Matter Expertise**: Acted as a subject matter expert in data engineering, providing guidance and education to team members and end-users on data products.

Python Developer Intern, NIT, Warangal, India   April 2019 -June 2019

- Developed a Python module facilitating student room allocation in a hostel comprising over 1.8k+ rooms.
- Acknowledged by NIT-Warangal management with an **appreciation letter** for significant contributions to the project.

Certifications

- Databricks Certified Associate Developer for Apache Spark 3.0 (In progress, August 2024)
- Databricks Generative AI Fundamentals
- AWS Certified Cloud Practitioner
- Astronomer Certification for Apache Airflow Fundamentals
- SQL Badge – LeetCode
- Google IT Automation with Python Specialization

Achievements

- Finalist in the 2024 Business Intelligence & Analytics Leadership Development Program Case Competition (BI&A LDP) among 400 students, presented an innovative AI architecture using AWS for the insurance sector, emphasizing efficient data extraction, ingestion, and analytical processing strategies.
- Secured two bronze medals in Hacker Coding Competitions.