

ROHITH GADE

+1(774)518-2092 | rohithgade0911@gmail.com | [LinkedIn Profile](#)

SUMMARY

Proactive Data Engineer with expertise in military-grade data analytics and secure data management. Skilled in Python, SQL, and NoSQL for complex data modeling, ETL/ELT processes, and cybersecurity-focused data wrangling.

TECHNICAL SKILLS

Languages: Python (pandas, PySpark, SQLAlchemy), SQL (MySQL, PostgreSQL, NoSQL), JSON, XML, R, Java.

Data Tools: AWS (S3, Lambda, Glue, Quicksight, Athena, IAM, Redshift), Azure, GCP, Git, Kafka, Splunk, Power BI, Tableau.

Big Data Ecosystem: Apache Kafka, Hadoop, Airflow, Hive, Spark, Informatica, Snowflake, Docker, Amazon EMR, Kubernetes.

Other Skills: Data Modeling, TDD, Data Warehouse, Machine Learning, Data Analysis, ETL/ELT, Agile methodologies, CI/CD Pipelines.

PROFESSIONAL EXPERIENCE

UNIVERSITY OF MASSACHUSETTS DARTMOUTH USA

Dartmouth, MA

Graduate Teaching Assistant (Advanced Mathematical Statistics)

September 2023 - May 2024

- I enhanced curriculum delivery and student engagement, guiding over 250 students through complex statistical methodologies such as linear and logistic models, clustering, bootstrap models, and cross-validation methods, which facilitated their academic success and application skills.

TATA CONSULTANCY SERVICES INDIA

Hyderabad, India

Associate Data Engineer

August 2021 - December 2022

- As a key contributor to a Scrum team, I enhanced CI/CD pipelines and service architecture for efficient data integration and processing, employing Apache Airflow for ETL/ELT workflows from AWS S3 to Snowflake and expertise in Python and SQL.
- Additionally, I fostered cross-functional collaboration on Safety Stock analysis through TDD in GitHub CI, leveraging Machine Learning in AWS SageMaker and developing ETL/ELT jobs with AWS Glue for Redshift, complemented by continuous learning in Splunk, ServiceNow, and Salesforce CRM, aligning with client needs and project goals.

ACADEMIC PROJECTS

UNIVERSITY OF MASSACHUSETTS DARTMOUTH

Advanced ETL Pipeline for Reddit Data Analysis and Warehousing

- Developed a robust ETL pipeline to facilitate the ETL of Reddit data into AWS Redshift for insightful analytics.
- Utilized a combination of Apache Airflow for workflow automation, Celery for task queuing, PostgreSQL for data storage, AWS services like S3 for data storage, Glue for data preparation, Athena for interactive queries, and Redshift for data warehousing.

Data Engineering YouTube Analysis Project

- Spearheaded the development of a data ingestion and ETL system for YouTube video analysis, leveraging AWS technologies like S3 for scalable storage, AWS IAM for secure access management, and AWS Glue for efficient data integration and transformation.
- Conducted comprehensive data analysis, employing AWS Athena for interactive data querying, extracting key insights from a Kaggle dataset of daily trending YouTube videos, and made visualizations using AWS Quicksight.

Data-Driven Smart City Optimization

- Engineered a data-driven smart city architecture using Docker for containerization, and AWS for cloud-based data processing and storage, alongside Apache Spark for analytics, enhancing urban mobility and safety by leveraging real-time IoT data integration and analysis for traffic and public safety improvements.
- Led the implementation of IoT integration using Apache Zookeeper, Kafka, and Spark, managed via AWS, and developed PowerBI visualization dashboards for dynamic urban management, improving response to emergent situations.

Next-Gen Data Workflow Management with Kubernetes and Airflow

- Engineered a scalable data engineering environment utilizing Kubernetes for orchestration and Apache Airflow for workflow management, demonstrating advanced skills in deploying robust data processing solutions.
- Implemented and optimized Directed Acyclic Graphs (DAGs) for efficient workflow scheduling and automation, showcasing my capability in enhancing data operations and process efficiency within a Kubernetes-driven architecture.

EDUCATION

University of Massachusetts Dartmouth

Massachusetts, USA

Masters in Data Science

May 2024

Jawaharlal Nehru Technological University Hyderabad

Hyderabad, India

Bachelors in computer science engineering

June 2021