

SAI KIRAN MAMIDI

+14057191238 | saireddy3795@gmail.com | linkedin.com/in/saikiranmamidi/

SUMMARY

Data Engineer with over 5 years of experience in data management, including data integration, modeling, and quality. Skilled in developing data strategies aligned with business objectives and defining data architecture principles. Proficient in Spark, Python, SQL, ETL, and data modeling, with a strong foundation in collaborating with stakeholders and IT teams.

Technical Skills

- Programming: Python, Java, SQL, Scala, C, C++
- Database/Frameworks: MS SQL, MySQL, Oracle, PL/SQL, Cosmos DB, PostgreSQL, NoSQL, MongoDB, Big Query, Cassandra, DynamoDB, Microservices, Airflow, Hadoop, Hive, Map Reduce
- Tools: Power BI, Tableau, Grafana, Terraform, Github, Confluence, Jenkins, Jira, Azure DevOps, Kubernetes, Docker
- Big Data Tools: Spark, Kafka, Snowflake, Databricks,
- Cloud Services: AZURE (Databricks, ADF, Data Lake, Synapse, Cosmos DB, Event hubs), AWS (EC2, S3, EMR, Redshift, Glue, Kafka, Lambda, SNS, SQS, Kinesis), GCP(DataProc, GCS, Dataplex)

PROFESSIONAL EXPERIENCE

Altria

Richmond, VA, USA

Data Engineer

December 2021 - Present

- Developed a robust data platform using lake house architecture in Databricks, integrating batch and real-time data processing. Orchestrated via Azure Data Factory and Azure SQL database.
- Designed a flexible data ingestion layer, enabling efficient integration of over 150 datasets, significantly reducing development costs and expediting data availability.
- Engineered and maintained scalable ETL pipelines processing approximately 1 TB of data daily, with a focus on data quality and adherence to complex business rules.
- Optimized data processing by refining Spark jobs and SQL queries, implementing partitioning and data compression techniques, resulting in a 30% cost reduction.
- Created a data warehouse employing a star schema, optimizing table design and optimizing the table reads.
- Led the design and implementation of an MLOps architecture, enabling efficient productionization of machine learning models using MLFlow.
- Mentored junior data engineers, imparting best practices and technical expertise in data engineering.

Microsoft

San Francisco, CA, USA

Data Engineer

September 2021 - December 2021

- Analyzed data collected during AB testing processes and identified abnormal patterns, providing valuable insights for decision-making.
- Provided technical support and troubleshooting assistance to Microsoft's clients on Cosmos DB-related issues, ensuring smooth operations and customer satisfaction.
- Built and maintained ETL pipelines, ensuring data integrity, reliability, and optimal performance throughout the data processing workflows

Standard and Poors

Hyderabad, Telangana, India

Data Engineer

July 2016 - July 2019

- Architected an automated ETL data pipeline using AWS services, enhancing data integration and unification from diverse sources.
- Employed Spark Streaming with Kafka to execute real-time data transformations, achieving a consolidated learner data model via delta tables.
- Optimized SQL queries and backend processes, significantly enhancing load times and efficiency across various schemas.
- Developed a high-performance Java API, handling over 1 million daily requests, facilitating complex data transformations and predictive analytics integration with Excel.
- Established robust CI/CD workflows with GIT and Jenkins, improving deployment efficiency and reliability in QA, UAT, and production environments.
- Collaborated across teams to enhance batch job efficiency by 50%, streamlining operations.

EDUCATION

University of Oklahoma

August 2019 - August 2021

Master's, Data Science

GPA: 3.75

JNTUH

September 2012 - May 2016

Bachelor's, Computer Science

GPA: 4