SHIVA RAMAN

Data Engineer

7397537266 • shiv.connect@outlook.com • Chennai, India

Summary

I am a Data Engineer with over 6.7 years of hands-on experience in developing sophisticated data solutions across cloud and on-premise environments. My expertise includes using tools such as PySpark, SQL, and Azure to enhance data processing efficiency and cut operational costs. I excel in ETL processes and possess a solid understanding of Azure components that optimize data pipeline reliability

Skills

Alteryx · Apache Spark · Azure · Azure Cloud · Azure Data Factory · Azure Data Lake · Azure Databricks · Databricks · Dataflow · ETL · ETL Pipelines · Git · GitHub · Jenkins · PySpark · Python · SQL · Shell Scripting

Experience

TCS (client – Wells Fargo)
Chennai, India

Azure Data Engineer
04/2023 - Present

A consultancy company specializing in technology, services, and solutions

- Migrated Alteryx workflows to a robust Azure Cloud Data Pipeline solution for Marketing Mix Modelling, focusing on historical and incremental data processing
- Managed import data from API URLs, Outlook reports, and SFTP to Azure ADLS Gen2 and export operations for marketing and sales data into Medallion Architecture in ADLS Data Lake Gen2 for efficient storage and processing
- Designed and deployed data pipelines with Azure Data Factory, and Azure Databricks for seamless data flow from the Raw Layer to Aggregate Layer & Reporting Layer
- Spearheaded migration of legacy systems to modern Azure cloud infrastructure, enhancing data processing capabilities through optimized Databricks solutions
- Engineered robust data transformation processes using PySpark and SQL, ensuring seamless data flow from diverse sources to Azure Data Lake Storage Gen2
- Led performance optimization initiatives for Azure Databricks jobs, resulting in improved processing efficiency and reduced operational costs

TCS (Client – PayPal)
Chennai, India

Databricks Developer/ Data Engineer
07/2021 - 03/2023

A global technology company that provides various services and solutions

- Engineered high-performance ETL pipelines using PySpark for On-prem Data Processing and Analytical Operations
- Implemented automated API integration workflows, transforming manual data operations into streamlined processes for enhanced cross-team efficiency
- Engineered robust ETL pipelines using PySpark, optimizing data transformation workflows that reduced processing time by 4 hours per team member daily
- Streamlined data integration processes through API automation, replacing manual operations and enhancing cross-functional team productivity
- Partnered with data science teams to establish efficient data preparation frameworks, ensuring seamless access to high-quality datasets
- Architected scalable data blending solutions, implementing automated workflows that significantly improved operational efficiency

Experience

Atos Syntel (Client – PayPal)
Chennai, India

Data Ops & Python Engineer
05/2018 - 07/2021

- Worked closely with data stakeholders, to understand the root cause of the recurring data issues and help to align the ETL pipeline for corrective action
- Developed automation solutions using Python, and shell script, and fed the error logs into Optimus – a machine-learning auto-remediation engine
- Optimized the data pipeline and reduced the failure resolution time by 90% through automation
- Reduced monthly incidents by 80% for the data pipeline errors, and quickly remediated through the Optimus Remediation engine
- Increased Stakeholders and business confidence in the Data Pipeline by 95%
- Recognized by PayPal Leadership Team for contribution to the Data Failure remediation

Education

Rajalakshmi Engineering College
Chennai, India

Bachelor's Degree in Computer Science Engineering
05/2014 - 04/2018

Key Achievements

⭐ **Leadership Recognition**
Recognized by PayPal Leadership Team for contribution to the Data Failure remediation