Inamdar Mo Sameer

saminamdar444@gmail.com | +91-9130839685 | linkedin.com/in/Sameer-Inamdar | github.com/Sameer20200

Summary

Data Engineer with 3.10 years of experience in IoT, Telecommunication, and Healthcare industry. Specializing in Big Data technologies and Data Analysis. Expertise in developing and maintaining data engineering applications using Apache Spark, Scala, Pyspark, and Cloud. Proficient in complex SQL development, data warehousing, data modeling, ETL processes, data ingestion, and processing data. Seeking to leverage expertise in a challenging environment.

Technical Skills

- **Languages:** Python, Pyspark, Scala, Java, Bash, SQL, Shell-Scripting, Linux

- **Data Warehouses:** Hive, Impala, Redshift,

- **Big Data:** Hadoop, Hive, Spark, Databricks, Snowflake

- **Databases:** MySQL, Oracle, PostgresSQL, SQL Server

- **Workflow Management:** Control-M, Airflow

- **Cloud:** AWS (S3, EMR, Lambda), Azure (ADF, Synapse, ADLS), GCP (Bigquery, Dataproc)

- **Control System and Documentation:** Git, Bitbucket, JIRA, Confluence

- **Tools:** Intellij, Eclipse, PyCharm, Splunk, Pagerduty

- **DevOps-Tools:** CI/CD, Docker, Jenkins, Git ,GitHub

Professional Experience

- **Citiustech** Oct 2024 – Present
  Sr. Software Engineer Pune, MH

  - Developed and maintained an ETL Pipeline for Insurance data ingestion and enrichment, utilizing API calls data to ingest into HDFS Storage
  - Processed and transformed Insurance data using Spark-Scala, applying business logic to enrich the data with reference field for data integration
  - Developed data models and optimized queries for Redshift and PostgreSQL.
  - Developed new build processes and end-to-end pipelines in Spark-Scala, incorporating comprehensive data quality controls and maintaining best practices.
  - Developed LCK Data Creation Model to freeze the HealthPlan Data with Submission Year.
  - Designed and Implemented Serverless Data pipeline for scalable and dynamic scaling of ETL workload integration.
  - Stored the enriched Patients data in an external impala table for downstream analysis and reporting.
  - Developed and deployed Power BI dashboards and reports to visualize key business metrics, enabling actionable insights for stakeholders.
  - Designed and optimized Azure Databricks notebooks to extract, transform and load data from Data Lake to Azure Synapse.

- **Cognizant Technology Solutions** Feb 2022 – Oct 2024
  Associate Pune, MH

  - Developed a Python-based rule engine to ingest monitoring IOT data from Zabbix and Prometheus APIs, classifying incidents dynamically based on severity (P1, P2, P3).
  - Automated alert routing to support teams using integrated PagerDuty and Microsoft Teams channels, improving incident response times.

- Orchestrated monitoring and alerting workflows using Apache Airflow and Kubernetes pods for scalable, event-driven operations.
- Migrated Data Pipelines from on premise to AWS using Spark, Python, Boto3, AWS Glue, Athena, S3.
- Implemented unit and integration tests for Airflow DAGs to achieve over 90% test coverage, improving reliability and reducing production incidents.

- **Schneider**   Nov 2020 – Feb 2021
  Associate   Pune, MH
  - Developed and maintained ETL pipelines using Azure, Python and SQL to process large volumes of structured and unstructured data.
  - Worked on database and schema design which helped in 30% reduction in delivery time and job failures.
  - Wrote complex SQL queries to aggregate, filter, and join large datasets across multiple tables, enabling efficient reporting and analytics for business teams.
  - Implemented data validation and quality checks to ensure accuracy and consistency across datasets.

Education

- **Savitribai Phule Pune University, Pune**   2018 – 2021
  Bachelor of Electrical Engineering   Pune, MH

Certifications

- AWS Certified Developer - Associate (DVA-C02)

- AWS Certified Cloud Practitioner (CLF-C02)

- Microsoft Certified: Azure Data Engineer Associate (DP-203)

- Microsoft Certified: Azure Data Fundamentals (DP-900)

- Google Analytics (GA4) Certification

Recognitions

- Cognizant cheers Award 2023 for excellent contribution towards project.

- 5 star Award from FEA for outstanding mentorship and guidance to rural students.

Projects

- **Real-Time Analytics Platform:** Built a real-time analytics dashboard using Kafka, Spark to visualize streaming data.

- **Vehicles Data:** Designed and deployed a Azure data pipeline with Azure functions to ingest the Data to ADLS storage.

- **ETL Automation Framework:** Developed a reusable ETL framework using Airflow and Docker for batch data processing.

Publications

- **Relay Configuration using PCM600:** Optimized control panel functionality by adding logic settings, simplifying relay logic operations for multiple substation equipment, and extending existing logic to handle various fault conditions.