ANUJ PRIYADARSHI

+91 7656814505 ⋄ Odisha, India ⋄ ✉ anujpriyadarshi07@gmail.com ⋄ in anujpriyadarshi07

SUMMARY

Data Engineer with 6 years of experience designing scalable ETL pipelines, optimizing data models, and automating data workflows across AWS and Azure. Leveraged distributed data processing systems to enable businesses in multiple domains to make data-driven decisions and gain valuable insights.

EDUCATION

B.Tech in Computer Science Engineering — 2015 - 2019
College Of Engineering and Technology, Bhubaneswar — *CGPA: 9.02*

SKILLS

Languages & Database	Python, MS SQL, MySQL, OracleDB, MongoDB
Cloud Services	AWS(Lambda, EMR, S3, SNS, Glue, Redshift), Azure(ADF, Databricks, Data Lake)
Big Data & Streaming	PySpark, Kafka, Kinesis, Hive, Data Warehousing
Tools & Scheduler	Airflow, Rundeck, Github, Jira, Confluence, Docker, Azure Devops

EXPERIENCE

Senior Data Engineer (SDE-2) — Jan 2023 - Present
Sigmoid Analytics — *Bangalore, Karnataka*

- Built a scalable ETL solution to process IoT streaming data with high performance and reliability. Accelerated **new client onboarding from 2 weeks to just 3 days** by implementing a metadata-driven code framework.
- Orchestrated parallel ETL executions using Step Functions, Lambda, leveraging EMR to process 20+ GB of daily data for efficient calculation of **MTTR and MTBF** across manufacturing plants.
- Optimized MySQL queries through **strategic indexing and partitioning**, reducing dashboard refresh time by 80% and automated maintenance tasks partition which reduced monitoring effort by 50%.
- Developed an on-prem ETL solution, integrating data from 10+ Kafka topics using Airflow, PySpark.
- **Technologies**: PySpark, Python, AWS EMR, RDS, Lambda, Step Functions, Kinesis, Docker, Kafka

Data Engineer (SDE-1) — Jun 2021 - Dec 2022
Sigmoid Analytics — *Bangalore, Karnataka*

- Designed and implemented an ETL pipeline, processing daily data of 40+ GBs to feed commercial operations dashboards. Refined queries using pivot transformations, reducing table join time by 90%.
- Restructured database model, **reduce persistent tables by 60%**, minimize data redundancy and latency.
- Developed a **generic code framework** to streamline the ETL pipeline that ingests engagement data from Facebook, Salesforce, and other APIs in the gold layer, reducing the time to onboard new data sources by 75%.
- Databricks Unity Catalog was used to enhance data reusability across 3+ teams and minimize redundancy.
- **Technologies**: PySpark, Python, Databricks, ADF

Data Engineer — Jul 2019 - Jun 2021
Infosys Limited — *Hyderabad, Telangana*

- Implemented a **scalable ETL pipeline** for more than 150 flows, which reduced operation costs by 30%.
- Integrated structured and semi-structured data from API sources, email URL (**reducing manual effort by 100%**), CSV, Excel, PDF, custom file formats using Python, Spark, Hive and SAP BODS.
- Automated script-based configuration file generation, reducing manual effort by 90% and improving efficiency.

AWARDS

Bravo Award - Sigmoid
Recognized twice for delivering a high-impact data pipeline that significantly improved business insights.

Spot Award - Infosys
Awarded for automating critical processes that reduced project timelines.