Raj Patel

Principal Data Engineer

raj.patel@devmail.com   +91-98765-43210   Bangalore, India

LinkedIn   GitHub   LeetCode   Portfolio

PROFESSIONAL SUMMARY

Data engineer with 6 years of experience building large-scale data platforms. Specialized in real-time streaming pipelines and ML infrastructure. Built systems processing 50TB daily.

TECHNICAL SKILLS

Languages: Python, Scala, Java, SQL

Frameworks: Apache Spark, Kafka, Airflow, Flink

Databases: Snowflake, PostgreSQL, Cassandra, Hive

Cloud: AWS, GCP

Tools: Docker, Kubernetes, MLflow, dbt

PROFESSIONAL EXPERIENCE

Lead Data Engineer
2021-04 - present

Flipkart

Leading data platform team of 8 engineers. Built real-time recommendation pipeline processing 100M+ events/day.

Technologies: Python, Scala, Apache Spark, Kafka, Airflow, Snowflake, AWS

Senior Data Engineer
2019-06 - 2021-03

Infosys

Designed ETL pipelines for Fortune 500 banking client. Reduced data processing time from 8 hours to 45 minutes.

Technologies: Python, Java, Apache Spark, Hive, HDFS, Airflow, GCP

Data Engineer
2018-07 - 2019-05

TCS

Built data warehouse for telecom client. Migrated legacy Oracle systems to Hadoop ecosystem.

Technologies: Python, SQL, Hadoop, Hive, Sqoop

PROJECTS

StreamETL (open-source)
2022-09 - 2023-02 · 6 months · GitHub

Framework for building streaming ETL pipelines with exactly-once semantics. Used by 50+ companies.

Technologies: Scala, Apache Spark, Kafka, Delta Lake

DataQuality Monitor (personal)
2023-05 - 2023-08 · 4 months · GitHub

Automated data quality monitoring tool with Slack/PagerDuty alerting integration.

Technologies: Python, Great Expectations, Airflow, PostgreSQL

EDUCATION

Masters in Data Science
IIT Bombay, 2018 | GPA: 3.9

CERTIFICATIONS

AWS Data Analytics
Specialty

Databricks Certified
Data Engineer
Professional