Rahul Dubey

Data Engineer

I am a professional Data Engineer having 2 years of experience in building data-Intensive applications for Data Analytics, Data Warehousing, Stream Ingestion, Cloud Migration and Dashboard Engineering. I am specialized in advanced Python and SQL.

WORK EXPERIENCE

Data + Software Engineer - Tegna Inc. — September 2021 – Present
Capgemini India, Insight & Data, Mumbai, MH

- Working in the Marketing Research Analytics domain.
- Implemented Snowflake, Star and Dimensional Schema data models for data warehousing and BI.
- Worked on developing Attribution Models for potential Sales Opportunity Forecasting.
- Developed Sales Summary Dashboards for booked impressions and top revenue partners, customers for sales performance over OTT/Linear TV.
- Built and deployed data stream ingestion application for ingesting transactions data from Ad-Servers using Python, SQL, Dataflow, Cloud Pub/Sub, GKE, Airflow, Big Query, Snowflake
- Developed Microservice-based Backend System for placing Ad-Creatives over SSP platform from Mendix Application Front-end using GCP Cloud Functions, Cloud Pub/Sub, Cloud Run, Cloud Scheduler, BigQuery.
- Built and deployed Python Microservice-based backend system for Angular Application using GCP Cloud Functions, Pub/Sub, Cloud Run, Celery, Redis, Snowflake, Github, CI/CD and Cloud Build.

Software Engineer - NBCU & Univision — June 2022 – August 2022
Capgemini India, Insight & Data, Mumbai, MH

- Worked with the Center-Of-Excellence (CoE) team for Cloud Migration and Data Governance project conducted in collaboration with Univision and NBCU.
- Developed data integration pipelines for both Batch and Streaming sources using GCP Composer, Dataflow, BigQuery, Pub/Sub.
- Built data validation automation solutions for Data Quality Checks and Governance using Python, SQL.
- Implemented Cybersecurity Breach Detection solution using Python.

ML Research Assistant - Health Informatics — April 2021 – July 2021
University of Glasgow, Glasgow, Scotland, UK

- Worked as ML Research Assistant under renowned for the Department of Health Informatics at University of Glasgow.
- Developed base-line models for Human Motion Gait Analysis.
- Designed and presented Deep Learning Analysis and Solutions for inferring neuro-pathological gait to the department of health informatics at University of Edinburgh and University of Oxford.

CONTACT

- Mumbai, MH
- +91-8305573987
- rahuldubey391@gmail.com
- linkedin.com/in/rahuldubs123
- github.com/RahulDubey391

SKILLS

Programming Languages:
- Python, Go, SQL, PL/SQL

Big Data/Storage:
- PySpark, Dask, Ray, Pandas, Numpy
- Snowflake, BigQuery, Redshift
- Apache Airflow, Cloud Scheduler
- HDFS, GCS, S3, Firestore

Software Engineering:
- OOPs/Procedural/Functional
- Multiprocessing/Multithreading
- Asyncio, Event-Loops
- Task Queues - Celery, Redis
- Message Queue - Kafka, Cloud Pub/Sub

API & Cloud Hosting:
- Google Cloud App Engine
- Cloud Run, Cloud Function

CI/CD & Container Orchestration:
- Kubernetes
- Docker
- Git, Github

BI Tools & Data Visualization:
- ThoughtSpot, GDS-360, Looker
- Redash, Plotly, Matplotlib

Data Analytics :
- A/B Testing, T-Test, ANOVA,
- Modeling - Star, Snowflake, Dimension Schema

EDUCATION

University of Glasgow
Master of Science
Data Science
Glasgow City, Scotland, UK — March 2021

GGITS - RGPV
Bachelor of Engineering
Computer Science & Engineering
Jabalpur, MP, India — June 2018