DEEPAK RAO

DATA ENGINEER

CONTACT ME AT

- 68/7 Nagappa Building, East Mumbai-Pune Road, Dehuroad,Pune-412101
- drao11114@gmail.com
- www.linkedin.com/in/deepak-rao-023442205
- +91 8408944384

SKILLS SUMMARY

- Quick Learner
- Communication
- Team Player
- Positive Attitude

HOBBIES

- Football
- Cricket
- Basketball
- Swimming
- Console Games

AWARDS RECEIVED

- Shout out from senior manager and delivery manager for the ability to work under pressure.
- Appreciation Card for being identified as a flexible, hard worker, and smart worker.
- Excellence Award in the category of Passion and Customer Service.
- Client Appreciation in the form of a Differentiator Award

CERTIFICATIONS

- Google Cloud Certified Professional Data Engineer

SUMMARY

Data Engineer specializing in ETL Batch and Streaming Pipeline development with all stages of the development cycle. Well-versed in programming languages and frameworks including Apache Hadoop, Spark, Scala, Hive, Google Cloud Platform, Dataflow, Dataproc, IAM, GCS, BigQuery, PubSub, Cloud SQL, Airflow, Compute Instance, Unix, and Shell Scripting, Java, and Python.

WORK EXPERIENCE

Data Engineer, Bitwise Inc July 2018 - June 2021, Capgemini June 2021 - Present

- Implemented batch processes for loading data from the different file formats (fixed-width, Delimited, XLSX) to Hive - On-premises Hadoop Cluster.
- Migrated multiple on-premise processes into Google Cloud Platform ensuring the processes should be scalable, automated, exact replica of the legacy processes and should produce similar results as the legacy system.
- Involved in automating Batch and streaming data-load pipelines to load data from files(fixed-width and delimited) to BigQuery and to create workflows on GCP that are scheduled through Airflow.
- Involved in migration of Netezza System to Google Cloud Platform.
- Involved in Architecture, design, and build of ETL pipelines that load data from source files to Google Cloud Bigquery using Google Cloud Dataflow (Apache Beam).
- Implemented Upsert logic, Complex SQL transformations(Conditional Joins, lookups, case-when, etc) through Google Cloud Dataflow (Apache Beam).
- Migrated batch jobs from On-premises Hadoop cluster to Google Cloud Dataproc which is a managed service to run Apache Spark Jobs in Google Cloud.

EDUCATIONAL HISTORY

- 2014-2018 | Savitribai Phule Pune University | Dr. DY Patil Institute of Engineering Management and Research
- 2012-2014 |Maharashtra Board| Nowrosjee Wadia College