

AMRIT RAJ

☎ +918639027689 ✉ amritkarndit09@gmail.com 🔗 [linkedin.com/in/amrit-raj-805311162](https://www.linkedin.com/in/amrit-raj-805311162) 🏠 Pune, Maharashtra

Summary

Outcome-oriented Data Engineer specializing in Hadoop and Google Cloud Platform with dual certifications. Experienced in designing and enhancing data pipelines for on-premises and cloud setups. Proficient in GCP, Data Engineering, Data Warehousing and ETL, with robust collaborative and communication skill

Technical Skills

Tools/Languages: Python, Scala, SQL, BitBucket, Jenkins

Apache Ecosystem: Apache Hadoop, Apache NiFi, Apache Airflow, Apache Beam, Apache Hive, HDFS, Apache Kafka

Cloud Services: GCP, BigQuery, Data Flow, Cloud function, Cloud Composer, Cloud Storage, GKE, Pub/Sub

Software Development: Agile Methodology, Scrum, JIRA

Technical Experience

Accenture

Oct 2021 – May 2024

GCP Data Engineer

Pune, Maharashtra

- Built Airflow DAGs scheduled via Cloud Composer to orchestrate batch data jobs and streamline other processing tasks, which were deployed using Terraform and version controlled using BitBucket.
- Created, set up, and managed data pipelines using Apache NiFi for instant data intake and transformation.
- Designed and developed data pipelines using Google Cloud Dataflow to process and analyze large volume of Streaming data, resulting in **improved efficiency increase of 15%**.
- Implemented optimizations to enhance pipeline performances and scalability, resulting in significant **cost saving of 10%**.
- Optimized SQL queries by conducting performance tuning, root cause analysis, validation of enrichment logic, and ensuring data quality, resulting in a **reduction of 20%** in execution time and alleviating slot contention challenges within BigQuery.
- Implemented cloud infrastructure components such as Tables, Views, DAGs, Dataflow Jobs, and Cloud Functions using Terraform. Orchestrated deployment through CI/CD pipelines via Jenkins.
- Utilized Cloud Monitoring Metrics, user defined metrics and Cloud Logging to evaluate and monitor performance of various jobs, promptly alerting in the event of job failures or abnormalities.
- Collaborated with cross-functional teams to streamline SQL query processes, resulting in a **15% increase** in ETL process efficiency.
- Proficiency in Python extends to the ability to manage a variety of task within the language, including but not limited to coding, scripting, automation and data analysis.
- Actively up-skilled in various Cloud Services and Big Data Processing tools like Spark and Data Proc through various Client based POC's and hands-on experience.
- Utilize JIRA for comprehensive backlog management, progress tracking and ensuring transparency across the development cycle.

Outdoor Route Optimization | Tata Consultancy Service

June 2024 - Present

- Led end-to-end development of data pipelines utilizing Google Cloud Platform (GCP), Apache NiFi, Cloud Composer, and Cloud Dataflow.
- Conducted performance analysis and continuous optimization to ensure the scalability and cost-effectiveness of data processing solutions.

Certifications

Google Cloud Certified Professional Data Engineer

Google Cloud Certified Associate Cloud Engineer

Education

DIT University

June 2017 – June 2021

Bachelor of Technology in Computer Science & Engineering

Dehradun, Uttarakhand